Gradient formula for the beta function of 2D quantum field theory

2010 J. Phys. A: Math. Theor. 43 215401

(http://iopscience.iop.org/1751-8121/43/21/215401)

J. Phys. A: Math. Theor. **43** (2010) 215401 (28pp) doi:10.1088/1751-8113/43/21/215401

Gradient formula for the beta function of 2D quantum field theory

Daniel Friedan[1,2] **and Anatoly Konechny**[3,4]

[1] Department of Physics and Astronomy Rutgers, The State University of New Jersey, Piscataway, NJ 08854-8019, USA
[2] Natural Science Institute, The University of Iceland, Reykjavik, Iceland
[3] Department of Mathematics, Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS, UK
[4] Maxwell Institute for Mathematical Sciences, Edinburgh, UK

E-mail: friedan@physics.rutgers.edu and anatolyk@ma.hw.ac.uk

Received 23 February 2010
Published 4 May 2010
Online at stacks.iop.org/JPhysA/43/215401

Abstract
We give a non-perturbative proof of a gradient formula for beta functions of two-dimensional quantum field theories. The gradient formula has the form $\partial_i c = -(g_{ij} + \Delta g_{ij} + b_{ij})\beta^j$ where β^j are the beta functions, c and g_{ij} are the Zamolodchikov c-function and metric respectively, b_{ij} is an antisymmetric tensor introduced by Osborn and Δg_{ij} is a certain metric correction. The formula is derived under the assumption of stress–energy conservation and certain conditions on the infrared behavior the most significant of which is the condition that the large-distance limit of the field theory does not exhibit spontaneously broken global conformal symmetry. Being specialized to nonlinear sigma models this formula implies a one-to-one correspondence between renormalization group fixed points and critical points of c.

PACS numbers: 04.60.Kz, 11.10.Gh, 11.10.Hi, 11.25.Sq

1. Introduction

Change of scale in quantum field theories (QFTs) is governed by renormalization group (RG) transformations. If a space of theories is parameterized by coupling constants $\{\lambda^i\}$, the RG transformations are governed by a beta function vector field:

$$\mu\frac{\mathrm{d}\lambda^i}{\mathrm{d}\mu} = \beta^i(\lambda). \tag{1.1}$$

The idea that RG flows could be gradient flows, that is

$$\beta^i(\lambda) = -G^{ij}(\lambda)\frac{\partial S(\lambda)}{\partial\lambda^j} \tag{1.2}$$

for some metric $G^{ij}(\lambda)$ and potential function $S(\lambda)$ defined on the theory space, has some history. One of the earliest papers devoted to this question was [2]. It was suggested in that paper that RG flows are gradient flows in a wide variety of situations. Gradient flows have some special properties. Thus, if the metric G^{ij} is positive definite, the scale derivative of the potential function is negative definite:

$$\mu\frac{\mathrm{d}S}{\mathrm{d}\mu} = \beta^i\frac{\partial S}{\partial\lambda^i} = -G_{ij}\beta^i\beta^j \leqslant 0 \tag{1.3}$$

and therefore S monotonically decreases along the flow. This demonstrates irreversibility of the RG flows and forbids limiting cycle behavior. Another appealing property of gradient flows is that the matrix of anomalous dimensions $\partial_i\beta^j$ is symmetric and thus their eigenvalues at critical points, that give critical exponents, are always real.

The first perturbative computations in support of this idea were done for four-dimensional theories [3]. Later more evidence was found in the context of two-dimensional general sigma models [4, 5]. In [5] a gradient formula of the form (1.2) was formulated for such models and shown to hold up to two loops for a particular class of sigma models. A crucial ingredient for a gradient formula for general sigma models was the introduction of the dilaton field [6, 7]. It was shown in [8, 10] that including the dilaton couplings into a general sigma model one finds that the vanishing beta function equations are equivalent to critical points of a certain functional at the leading order in α'. A gradient formula of the form (1.2) was checked for general sigma models in [9] to the first two orders in α'. In string theory conformal sigma models describe strings propagating on the sigma model target manifolds. The sigma model couplings parameterize a metric G_{IJ}, an antisymmetric tensor B_{IJ} and a dilaton field Φ defined on the target space manifold. The gradient property (1.2) attains a special significance in this context becoming a manifestation of the string action principle. The condition for conformal invariance is that the beta functions vanish: $\beta^G = \beta^B = \beta^\Phi = 0$. It is equivalent to string equations of motion. The gradient property (1.2) thus means that the string equations of motion arise by varying a functional of couplings—S, which can be identified with the string action functional.

Another reinforcement of the gradient conjecture (1.2) for two-dimensional theories came from the Zamolodchikov c-theorem [11]. The last one is a general theorem applicable to unitary 2D theories that states that there is a function c on the space of theories that monotonically decreases along the RG flows and coincides with the Virasoro central charge at fixed points. (We give a slightly modified proof of this theorem in section 3.) The theorem was proved by constructing c whose scale derivative takes the form of the right-hand side of (1.3) with a certain positive definite metric. It was natural to conjecture that a gradient formula of the form (1.2) holds with S being the c function and G_{ij}—the Zamolodchikov metric. This was shown to hold at the leading order in conformal perturbation theory near fixed points [11, 13]. In the context of nonlinear sigma models this idea was discussed in [15]. It was argued in [15] that for the purposes of string theory the c-function cannot provide a suitable potential function (we comment more on this in section 10). Other potential functions for RG flows of nonlinear sigma models were considered in [12, 14, 15, 17, 18] which were shown to be related to the central charge and to each other. In [18] a potential function for nonlinear sigma models was constructed assuming the existence of a sigma model zero mode integration measure with certain properties. It was shown that a measure with the required properties can be constructed infinitesimally but a proof of the integrability of that construction is still lacking. An essential tool proposed in [18] for deriving gradient formulas was the use of Wess–Zumino consistency conditions on local Weyl transformations in the presence of curved metric and sources. This technique was applied in [19] to a class of quantum field theories subject to certain power counting restrictions. It was shown that for these theories a gradient formula holds which is of a slightly different form than (1.2):

$$\partial_i c = -g_{ij}\beta^j - b_{ij}\beta^j \tag{1.4}$$

where c and g_{ij} are Zamolodchikov's metric and c-function respectively [11] and b_{ij} is a certain antisymmetric tensor. The necessity to introduce an antisymmetric tensor along with Zamolodchikov's metric can be demonstrated by the use of conformal perturbation theory. Thus, it was shown in [23] by explicit perturbative calculations that the one-form $g_{ij}\partial^j c$ is not closed for some flows[5]. Still, as we will explain in the next section, Osborn's gradient formula (1.4), although very inspiring, falls short of providing a general gradient formula. The main content of the present work is a derivation of a gradient formula that generalizes formula (1.4) to a much wider class of theories that includes nonlinear sigma models as well.

To finish the historical overview we mention here that a general gradient formula was proven for boundary renormalization group flows in two dimensions [22]. Such flows happen in QFTs defined on a half plane (or a cylinder) when the bulk theory is conformal but the boundary condition breaks the conformal invariance. One of the implications of the boundary gradient formula is a proof of Affleck and Ludwig's g-theorem [21] which is a statement analogous to Zamolodchikov's c-theorem. A string theory interpretation of this gradient formula is that it provides an off-shell action for open strings. The boundary gradient formula was proved under certain assumptions on the UV behavior which are reminiscent of the power counting restrictions of [19]. Nevertheless, we will show in the present paper that any assumption of this kind can be dispensed with in proving a bulk gradient formula.

The paper is organized as follows. In section 2 after introducing some notations we explain in more detail Osborn's gradient formula (1.4) and the assumptions that went into proving it. We then state our main result—a general gradient formula (2.13)—and discuss the assumptions needed to prove it. In section 3 we give a proof of Zamolodchikov's formula and recast it in the form that we use as a starting point for proving the gradient formula. In section 4 the first steps of the proof are explained. At the end of those steps we express the quantity $\partial_i c + g_{ij}\beta^j + b_{ij}\beta^j$ built from the elements present in (1.4) via three-point functions with a certain contact operator present in them. To analyze these three-point functions we develop a sources and operations formalism in section 5. A short summary of the formalism is provided in section 5.2. After discussing the Callan–Symanzik equations in section 6 we resume the proof in section 7 putting to use the Wess–Zumino consistency conditions on the local renormalization operation and our infrared assumptions. At the end of section 7 an infrared regulated gradient formula is obtained. In section 8 the proof is concluded by removing the infrared cutoff. Section 9 contains a discussion of the properties of the gradient formula and the assumptions used in proving it. In section 9.5 the gradient formula is specialized to the nonlinear sigma model case and a proof is given of the correspondence between RG fixed points and stationary points of c. In section 10 we conclude with some final remarks.

2. The general gradient formula

In this paper we consider two-dimensional Euclidean quantum field theories equipped with a conserved stress–energy tensor $T_{\mu\nu}(x)$. The stress–energy tensor measures the response of the theory to metric perturbations, so that if $Z[g_{\mu\nu}]$ is a partition function defined on a two-dimensional plane with metric $g_{\mu\nu}(x) = \delta_{\mu\nu} + \delta g_{\mu\nu}$,

$$\delta \ln Z = \frac{1}{2}\iint \mathrm{d}^2x\,\langle \delta g_{\mu\nu} T^{\mu\nu}(x)\rangle. \tag{2.1}$$

[5] The obstruction to closedness occurs at the next-to-leading order in perturbation.

In two dimensions any metric can be made conformally flat so that $g_{\mu\nu}(x) = \mu^2(x)\delta_{\mu\nu}$ where the function $\mu(x)$ sets the local scale. A change of local scale is generated by the trace of stress–energy tensor $\Theta(x) \equiv g^{\mu\nu}T_{\mu\nu}(x)$:

$$\mu(x)\frac{\delta \ln Z}{\delta\mu(x)} = \langle\Theta(x)\rangle. \tag{2.2}$$

For correlation functions computed on $\mathbb{R}^2$ with constant scale μ the change of scale is obtained by integrating over an insertion of $\Theta(x)$:

$$\mu\frac{\partial}{\partial\mu}\langle\mathcal{O}_1(x_1), \ldots, \mathcal{O}_n(x_n)\rangle_c = \int \mathrm{d}^2x \langle\Theta(x)\mathcal{O}_1(x_1), \ldots, \mathcal{O}_n(x_n)\rangle_c. \tag{2.3}$$

Here $\mathcal{O}_1, \ldots, \mathcal{O}_n$ are local operators and the subscript c at the correlator brackets marks connected correlators.

Assume that a family of renormalizable QFTs is parameterized by renormalized coupling constants λ^i, $i = 1, \ldots, N$. We assume that an action principle [1] is satisfied. This means that for each coupling λ^i there exists a local operator $\phi_i(x)$ such that for any set of local operators $\mathcal{O}_1, \ldots, \mathcal{O}_n$

$$\frac{\partial}{\partial\lambda^i}\langle\mathcal{O}_1(x_1), \ldots, \mathcal{O}_n(x_n)\rangle_c = \int \mathrm{d}^2x\, \langle\phi_i(x)\mathcal{O}_1(x_1), \ldots, \mathcal{O}_n(x_n)\rangle_c. \tag{2.4}$$

Note that the integrability of the integrand in (2.3) and (2.4) assumes the appropriate infrared behavior of the correlators.

Assume further that the couplings λ^i can be promoted to local sources $\lambda^i(x)$ for the fields $\phi_i(x)$. The generating functional $\ln Z$ then in general depends on the scale factor $\mu(x)$ and the sources λ^i, and the action principle (2.4) means that in addition to (2.1) we have

$$\frac{\delta \ln Z}{\delta\lambda^i(x)} = \langle\phi_i(x)\rangle. \tag{2.5}$$

A correlation function of the form

$$\langle\phi_{i_1}(x_1)\phi_{i_2}(x_2)\ldots\phi_{i_n}(x_2)\Theta(y_1)\Theta(y_2)\ldots\Theta(y_m)\rangle_c \tag{2.6}$$

evaluated on a flat $\mathbb{R}^2$ can be obtained by taking variational derivatives of $\ln Z$ with respect to the sources λ^i and the metric scale factor μ and then setting the sources and the scale to be constant. In a renormalized theory the correlators (2.6) are distributions. They form a basic set of local physical quantities defined in a given QFT.

In a renormalizable QFT a change of scale can be compensated by changing the couplings λ^i according to (1.1). By the action principle (2.4) this implies that $\Theta(x) = \beta^i\phi_i(x)$ where β^i are the beta functions. This equation should be understood as an operator equation, that is, as an equation that holds inside correlation functions (2.6) up to contact terms (i.e. up to distributions supported on subsets of measure zero). The use of sources $\lambda^i(x)$ and non-constant Weyl factor $\mu(x)$ facilitates bookkeeping of the contact terms. In the presence of non-constant $\lambda^i(x)$ and $\mu(x)$ one can expand the difference $\Theta(x) - \beta^i(\lambda(x))\phi_i(x)$ in terms of derivatives of the sources and metric [19]. The expansion must by covariant with respect to changes of coordinates. This requirement ensures that the contact terms respect the conservation of stress–energy tensor. In [19] Osborn assumed that this expansion has the form

$$\Theta(x) - \beta^i\phi_i(x) = \tfrac{1}{2}\mu^2R_2(x)C(\lambda) + \partial^\mu[W_i(\lambda)\partial_\mu\lambda^i] + \tfrac{1}{2}\partial_\mu\lambda^i\partial^\mu\lambda^jG_{ij}(\lambda) \tag{2.7}$$

where

$$\mu^2R_2(x) = -2\partial_\mu\partial^\mu \ln\mu(x) \tag{2.8}$$

is the two-dimensional curvature density. Note that in (2.7) C, W_i and G_{ij} are functions of λ evaluated on $\lambda^i(x)$ that depend on x via $\lambda^i(x)$ only. Effectively equation (2.7) gives a local version of the renormalization group equation. Using the Wess–Zumino consistency conditions for the local renormalization group transformations (2.7) Osborn derived a gradient formula [19]

$$\partial_i c + g_{ij}\beta^j + b_{ij}\beta^j = 0 \tag{2.9}$$

where c and g_{ij} are the Zamolodchikov c-function and metric respectively[11] defined in terms of two-point functions as

$$c = 4\pi^2\left(x^\mu x^\nu x^\alpha x^\beta - x^2 g^{\mu\nu}x^\alpha x^\beta - \tfrac{1}{2}x^2 x^\mu g^{\nu\alpha}x^\beta\right)\langle T_{\mu\nu}(x)T_{\alpha\beta}(0)\rangle_{c/\Lambda|x|=1} \tag{2.10}$$

$$g_{ij} = 6\pi^2\Lambda^{-4}\langle\phi_i(x)\phi_j(0)\rangle_{c/\Lambda|x|=1} \tag{2.11}$$

where Λ^{-1} is a fixed arbitrary 2D distance. The tensor b_{ij} is an antisymmetric two-form that can be expressed as

$$b_{ij} = \partial_i w_j - \partial_j w_i, \qquad w_i = 3\pi\int \mathrm{d}^2x\, x^2\theta(1-\Lambda|x|)\langle\phi_i(x)\Theta(0)\rangle_c \tag{2.12}$$

where Λ is the same mass scale used in the definitions of c and g_{ij}. The most restrictive assumption in [19] appears to be the form of expansion (2.7). The fact that the expansion does not go beyond the second order in derivatives suggests a certain power counting principle. Such a principle could be provided in the vicinity of an ultraviolet fixed point by the standard power counting arguments for renormalizability. Even with such a counting principle expansion (2.7) is too restrictive. Thus, it omits terms of the form $\partial_\mu\lambda^i J_i^\mu(x)$ where $J_i^\mu(x)$ are local vector fields which can be prescribed engineering dimension 1. Such terms in the scale anomaly can be generated by near marginal perturbations near fixed points. In particular they are present in generic current–current perturbations of Wess–Zumino–Witten theories [27]. Another class of theories for which (2.7) is too restrictive is general nonlinear sigma models. In this case one needs to allow the quantities C, W_i and G_{ij} in (2.7) to have a non-trivial operator content. The case of sigma models was covered separately in [19] (see also [12, 14, 15, 17, 18] and references therein). It was shown that a gradient formula analogous to (2.9) can be derived provided a sigma model integration measure with certain properties exists. In the present paper we will go beyond Osborn's UV assumptions allowing for an arbitrary local covariant expansion with operator-valued coefficients replacing (2.7). Making instead assumptions about the infrared behavior we derive a general formula

$$\partial_i c + (g_{ij} + \Delta g_{ij})\beta^j + b_{ij}\beta^j = 0. \tag{2.13}$$

The metric correction Δg_{ij} is constructed via two-point functions of ϕ_i with the currents $J_j^\mu(x)$ arising from the expansion generalizing expansion (2.7) (see formulas (8.2) and (8.15)). Alternatively Δg_{ij} can be expressed via three-point functions with the pure-contact field $D(x) = \Theta(x) - \beta(x)$ (formula (8.3)). Formula (2.13) is derived under two separate assumptions on the infrared behavior. The first assumption is that the action principle (2.4) holds for one- and two-point functions of operators ϕ_i that assumes that these functions are at least once differentiable. This ensures in particular that the c-function is once differentiable. The second assumption is that for any vector field $J_\mu(x)$ we have

$$\lim_{|x|\to\infty}|x|^3\langle J_\mu(x)T_{\alpha\beta}\rangle_c = 0. \tag{2.14}$$

This condition is equivalent to requiring that the long-distance limit of the QFT does not exhibit spontaneously broken global conformal symmetry. (Recall that at fixed points special conformal symmetry requires $T(z)$ to decay at infinity as $|z|^{-4}$.) As a simple example in section 9.2 demonstrates, this condition is essential. If in a scale invariant theory the global conformal symmetry is broken via boundary conditions at infinity, the value of the central charge may vary with moduli.

Our considerations include the nonlinear sigma model case. We thus show that in order to have a gradient formula we may replace the somewhat obscure technical assumption on the measure given in [18] by a more conceptually clear assumption on the stress–energy tensor behavior (2.14) which we show to be a necessary assumption in section 9.2. A question remains, of course, how one can check whether our infrared conditions hold in any given theory. Since in the nonlinear sigma model the expectation values of diffeomorphism invariant local operators are believed to be free of perturbative infrared divergences, they must be analytic in the couplings ([24, 25]). This means that the first infrared assumption can be controlled in perturbation theory. It is less clear to us whether one can control the infrared behavior of $T_{\mu\nu}$ perturbatively. We are planning to discuss applications of our general result (2.13) to nonlinear sigma models in more detail in a separate paper [27].

3. Zamolodchikov's formula

Zamolodchikov proved in [11] the following formula:

$$\mu\frac{\partial c}{\partial\mu} = -\beta^i g_{ij}\beta^j \tag{3.1}$$

where μ is the RG scale, c is the c-function (2.10) and g_{ij} is the metric introduced in (2.11). This formula implies that c decreases under the renormalization group flow and is stationary exactly at the fixed points. c is normalized so that at fixed points its value coincides with the value of the Virasoro central charge.

Note that the c-function and the metric g_{ij} depend on Λ only through the dimensionless ratio Λ/μ because according to (2.1) and (2.4) the fields $T_{\mu\nu}(x)$ and $\phi_i(x)$ are densities in x, implying that their two-point functions take the form

$$\begin{aligned}
&\langle T_{\mu\nu}(x)T_{\alpha\beta}(0)\rangle_c = \mu^4 F_{\mu\nu\alpha\beta}(\mu x),\\
&\langle \phi_i(x)\phi_j(0)\rangle_c = \mu^4 F_{ij}(\mu x),\\
&\langle T_{\mu\nu}(x)\phi_i(0)\rangle_c = \mu^4 F_{\mu\nu,i}(\mu x).
\end{aligned} \tag{3.2}$$

Before we set out to prove the general gradient formula it is instructive to go over a proof of formula (3.1). One way to prove equation (3.1) is to derive alternative formulas for c and g_{ij}

$$c = -\int d^2x\ G_\Lambda(x)\langle\Theta(x)\Theta(0)\rangle_c \tag{3.3}$$

$$g_{ij} = -\Lambda\frac{\partial}{\partial\Lambda}\int d^2x G_\Lambda(x)\langle\phi_i(x)\phi_j(0)\rangle_c \tag{3.4}$$

where

$$G_\Lambda(x) = 3\pi x^2\theta(1-\Lambda|x|). \tag{3.5}$$

These are the formulas for c and g_{ij} that we will use in the proof of the gradient formula. Equation (3.1) follows immediately from formulas (3.3) and (3.4):

$$\begin{aligned}
\mu\frac{\partial c}{\partial\mu} &= -\Lambda\frac{\partial c}{\partial\Lambda} = \Lambda\frac{\partial}{\partial\Lambda}\int d^2x\ G_\Lambda(x)\langle\Theta(x)\Theta(0)\rangle_c\\
&= \int d^2x\ \Lambda\frac{\partial G_\Lambda(x)}{\partial\Lambda}\langle\beta^i\phi_i(x)\beta^j\phi_j(0)\rangle_c\\
&= -\beta^i g_{ij}\beta^j.
\end{aligned} \tag{3.6}$$

Replacing $\langle\Theta(x)\,\Theta(0)\rangle_c$ by $\langle\beta^i\phi_i(x)\,\beta^j\phi_j(0)\rangle_c$ in the second line is allowed because they differ only by a contact term in x, which gives no contribution since the smearing function $\Lambda\partial G_\Lambda(x)/\partial\Lambda$ is supported away from $x=0$.

While formula (3.4) is evidently equivalent to formula (2.11) the equivalence of formulas (2.10) and (3.3) for c is shown as follows. Combine the special identity in two spacetime dimensions

$$(x^2g^{\mu\nu}g^{\alpha\beta}-g^{\mu\nu}x^\alpha x^\beta-x^\mu x^\nu g^{\alpha\beta}+2g^{\mu\alpha}x^\nu x^\beta-x^2g^{\mu\alpha}g^{\nu\beta})\langle T_{\mu\nu}(x)T_{\alpha\beta}(0)\rangle_c=0 \qquad (3.7)$$

with the Ward identity

$$\partial^\mu\langle T_{\mu\nu}(x)T_{\alpha\beta}(0)\rangle_c=0 \qquad (3.8)$$

and CPT invariance

$$\langle T_{\mu\nu}(x)T_{\alpha\beta}(0)\rangle_c=\langle T_{\mu\nu}(-x)T_{\alpha\beta}(0)\rangle_c=\langle T_{\alpha\beta}(x)T_{\mu\nu}(0)\rangle_c \qquad (3.9)$$

to calculate

$$\partial^\mu[(2x^\nu x^\alpha x^\beta-2x^2x^\nu g^{\alpha\beta}-x^2g^{\nu\alpha}x^\beta)\langle T_{\mu\nu}(x)T_{\alpha\beta}(0)\rangle_c]=-3x^2\langle\Theta(x)\Theta(0)\rangle_c. \qquad (3.10)$$

It follows from (3.10) that

$$\begin{aligned}
&-\int d^2x\; G_\Lambda(x)\langle\Theta(x)\Theta(0)\rangle_c\\
&=\pi\int d^2x\theta(1-\Lambda|x|)\partial^\mu[(2x^\nu x^\alpha x^\beta-2x^2x^\nu g^{\alpha\beta}-x^2g^{\nu\alpha}x^\beta)\langle T_{\mu\nu}(x)T_{\alpha\beta}(0)\rangle_c]\\
&=\pi\int d^2x\delta(1-\Lambda|x|)|x|^{-2}x^\mu(2x^\nu x^\alpha x^\beta-2x^2x^\nu g^{\alpha\beta}-x^2g^{\nu\alpha}x^\beta)\langle T_{\mu\nu}(x)T_{\alpha\beta}(0)\rangle_c\\
&=2\pi^2(2x^\mu x^\nu x^\alpha x^\beta-x^2x^\mu x^\nu g^{\alpha\beta}-x^2g^{\mu\nu}x^\alpha x^\beta-x^2x^\mu g^{\nu\alpha}x^\beta)\langle T_{\mu\nu}(x)T_{\alpha\beta}(0)\rangle_{c/\Lambda|x|=1},
\end{aligned} \qquad (3.11)$$

which demonstrates the equivalence of (2.10) and (3.3).

4. The proof of the gradient formula (first steps)

We start by defining a one-form r_i by the equation

$$\partial_ic+g_{ij}\beta^j+b_{ij}\beta^j+r_i=0 \qquad (4.1)$$

and show that the remainder term r_i can be expressed in terms of correlation functions of $\Theta(x)$ and $\phi_i(x)$ with the pure-contact field $D(x)=\Theta(x)-\beta(x)$. Infrared behavior of the correlation functions will be an important issue, so we introduce an IR cutoff at $|x|=L\gg\Lambda^{-1}$ and keep track of the error terms. Our assumptions about IR behavior will be designed to ensure the vanishing of the IR error in the limit $L\to\infty$.

We start out by recasting $g_{ij}\beta^j$ as

$$g_{ij}\beta^j=6\pi^2\Lambda^{-4}\langle\phi_i(x)\phi_j(0)\beta^j\rangle_{/\Lambda|x|=1}=6\pi^2\Lambda^{-4}\langle\phi_i(x)\Theta(0)\rangle_{/\Lambda|x|=1} \qquad (4.2)$$

which is valid because $\beta^j\phi_j(0)$ differs from $\Theta(0)$ only by contact terms. This can be further rewritten as

$$g_{ij}\beta^j=-\Lambda\frac{\partial}{\partial\Lambda}\int d^2x\,G_\Lambda(x)\langle\phi_i(x)\Theta(0)\rangle_c=\mu\frac{\partial}{\partial\mu}\int d^2x\,G_\Lambda(x)\langle\phi_i(x)\Theta(0)\rangle_c \qquad (4.3)$$

where the scaling property (3.2) was used in the last step. Finally using (2.3) we obtain

$$g_{ij}\beta^j=\int d^2y\int d^2x\;G_\Lambda(x)\langle\Theta(y)\phi_i(x)\Theta(0)\rangle_c. \qquad (4.4)$$

Formula (4.4) is infrared safe but as we want to impose the IR cutoff systematically, we write instead

$$g_{ij}\beta^j + E_1 = \int_{|y|<L} \mathrm{d}^2 y \int \mathrm{d}^2 x \; G_\Lambda(x)\langle \Theta(y)\phi_i(x)\Theta(0)\rangle_c. \tag{4.5}$$

The Ward identity gives the error term

$$\begin{aligned} E_1 &= \int_{|y|<L} \mathrm{d}^2 y \; \partial^\mu \left[y^\nu \int \mathrm{d}^2 x \; G_\Lambda(x)\langle T_{\mu\nu}(y)\phi_i(x)\Theta(0)\rangle_c \right] \\ &= 2\pi y^\mu y^\nu \int \mathrm{d}^2 x \; G_\Lambda(x)\langle T_{\mu\nu}(y)\phi_i(x)\Theta(0)\rangle_{c/|y|=L} \end{aligned} \tag{4.6}$$

which certainly vanishes in the limit $L \to \infty$.

We next turn our attention to the derivative $\partial_i c$. Assuming that c can be differentiated with respect to the coupling constants λ^i, we can write using formula (3.3) for c and the action principle (2.4)

$$\partial_i c = -\int \mathrm{d}^2 y \int \mathrm{d}^2 x \; G_\Lambda(x)\langle \phi_i(y)\Theta(x)\Theta(0)\rangle_c. \tag{4.7}$$

Again, we regularize in the IR as

$$\partial_i^L c = -\int_{|y|<L} \mathrm{d}^2 y \int \mathrm{d}^2 x \; G_\Lambda(x)\langle \phi_i(y)\Theta(x)\Theta(0)\rangle_c. \tag{4.8}$$

Formulas (4.5) and (4.8) can be combined to obtain

$$\begin{aligned} \partial_i^L c + g_{ij}\beta^j + E_1 &= \int_{|y|<L} \mathrm{d}^2 y \int \mathrm{d}^2 x G_\Lambda(x)\langle \Theta(y)\phi_i(x)\Theta(0) - \phi_i(y)\Theta(x)\Theta(0)\rangle_c \\ &= \int_{|y|<L} \mathrm{d}^2 y \int \mathrm{d}^2 x G_\Lambda(x)\langle [\beta(y) + D(y)]\phi_i(x)\Theta(0) - \phi_i(y)[\beta(x) + D(x)]\Theta(0)\rangle_c \\ &= -b_{ij}^L\beta^j + \int_{|y|<L} \mathrm{d}^2 y \int \mathrm{d}^2 x G_\Lambda(x)\langle D(y)\phi_i(x)\Theta(0) - \phi_i(y)D(x)\Theta(0)\rangle_c \end{aligned} \tag{4.9}$$

where we have introduced the two-form b_{ij}^L:

$$b_{ij}^L = \int_{|y|<L} \mathrm{d}^2 y \int \mathrm{d}^2 x G_\Lambda(x)\langle \phi_i(y)\phi_j(x)\Theta(0) - \phi_j(y)\phi_i(x)\Theta(0)\rangle_c. \tag{4.10}$$

Equation (4.9) can be written as

$$\partial_i^L c + g_{ij}\beta^j + E_1 + b_{ij}^L\beta^j + r_i^L = 0 \tag{4.11}$$

with

$$r_i^L = \int_{|y|<L} \mathrm{d}^2 y \int \mathrm{d}^2 x G_\Lambda(x)\langle \phi_i(y)D(x)\Theta(0) - D(y)\phi_i(x)\Theta(0)\rangle_c. \tag{4.12}$$

Equations (4.11) and (4.12) are the main results of this section. We will later show that under our assumptions on the infrared behavior the limits

$$\partial_i c = \lim_{L\to\infty} \partial_i^L c, \qquad b_{ij} = \lim_{L\to\infty} b_{ij}^L \tag{4.13}$$

exist. The error term E_1 goes to zero as $L \to \infty$. The remainder term r_i^L is expressed via correlation functions involving the pure-contact field $D(x)$. In order to investigate this term we develop a sources and operations formalism for calculating correlation functions of $D(x)$.

5. Sources and operations

In this section we present a general formalism that allows computing correlation functions of pure-contact fields using functional differential operators acting on functionals of sources and metric. The general exposition is somewhat tedious so for the reader's convenience we present the most important ingredients necessary to understand the proof of the gradient formula in a separate subsection 5.2.

5.1. General formalism

So far we have introduced the fields $\phi_i(x)$ as operators conjugate to the coupling constants λ^i that parameterize a renormalizable 2D QFT. It will be convenient to assume that the set ϕ_i is complete in a given class of fields which we denote by $\mathcal{F}$. The class of fields can be a complete set of spin-zero relevant and near-marginal fields. We could define such fields without a reference to a particular fixed point by requiring that the corresponding coupling constant belongs to some family of renormalizable theories with finitely many couplings (there are finitely many couplings for which β^i is not identically zero). This will not work for the nonlinear sigma models, for which the set of couplings is infinite, but in that case we could talk about near-relevant and near-marginal couplings using the engineering scaling dimensions introduced via free fields. As yet another possibility we could assume that the set $\{\phi_i\}$ spans all spin-zero local fields and works with a Wilsonian RG. We will keep the class of fields $\mathcal{F}$ unspecified throughout this section assuming only that $\mathcal{F}$ is closed under RG the precise sense of which we will discuss below. In general a field $O(x)$ is defined via its distributional correlation functions with other fields. If $O(x) \in \mathcal{F}$, the completeness of $\{\phi_i\}$ means that there are unique coefficients O^i such that the field $O(x) - O^i\phi_i(x)$ has vanishing correlation functions with all fields from $\mathcal{F}$ inserted away from x. The field $O(x) - O^i\phi_i(x)$ is thus a pure-contact field, that is its correlation functions are distributions supported on a subset of measure zero in x. We can define *ordinary* fields $O(x)$ as fields for which the correlations of $O(x) - O^i\phi_i(x)$ are zero as distributions. This means that the distributional correlation functions of such fields are obtained from those of the fields $\phi_i(x)$ by contracting them with the appropriate coefficients O^i.

Whatever $\mathcal{F}$ we choose it is essential that the trace of stress–energy tensor can be expanded in these fields: $\Theta(x) = \beta^i(\lambda)\phi_i(x)$. It is worth noting that the set ϕ_i may include total derivative fields. Although the correlation functions are independent of the corresponding coupling constants, the beta functions may be non-trivial and total derivatives may thus contribute to $\Theta(x)$. Let us further introduce sources $\lambda^i(x)$ for all fields $\phi_i(x)$ so that the generating functional $\ln Z$ depends on these sources and the metric scale factor $\mu(x)$ with equations (2.2) and (2.5) satisfied. This means that $\phi_i(x)$ and $\Theta(x)$ are represented by functional derivatives

$$\phi_i(x) = \frac{\delta}{\delta\lambda^i(x)}, \qquad \Theta(x) = \mu(x)\frac{\delta}{\delta\mu(x)} \tag{5.1}$$

which we chose to denote by the same symbols. The action of these functional derivatives on $\ln Z$ generates distributional correlation functions (2.6). To facilitate the use of differential operators in computing correlation functions we introduce a shorthand notation

$$\rangle\rangle = \ln Z \tag{5.2}$$

$$\langle\langle = \text{restriction of functionals to constant sources and flat 2D metric} \tag{5.3}$$

so

$$\langle\langle\phi_{i_1}(x_1)\cdots\Theta(y_1)\cdots\rangle\rangle = \langle\phi_{i_1}(x_1)\cdots\Theta(y_1)\cdots\rangle_c \tag{5.4}$$

where, on the left-hand side, the $\phi_i(x)$ and $\Theta(x)$ are functional differential operators (5.1), while on the right-hand side they are fields.

Define *operations* $\mathcal{O}(x)$ to be first-order local differential operators acting on functionals of the sources and 2D metric. The word local here means that the coefficients of the functional derivatives in an operation given at x can depend only on the values of $\lambda(x)$, $\mu(x)$ and finitely many derivatives thereof. An ordinary field $O(x) = O^i(\lambda)\phi_i(x)$ is naturally assigned an operation $\mathcal{O}(x) = O^i(\lambda(x))\phi_i$. Operations of this form we will call ordinary. An arbitrary operation $\mathcal{O}(x)$ gives rise to an ordinary field denoted $\underline{\mathcal{O}}(x)$ via

$$\langle \underline{\mathcal{O}}(x)\phi_{i_1}(x_1)\cdots\Theta(y_1)\cdots\rangle_c = \langle\langle \mathcal{O}(x)\phi_{i_1}(x_1)\cdots\Theta(y_1)\cdots\rangle\rangle. \tag{5.5}$$

Although the above formula specifies distributional correlation functions containing only a single $\underline{\mathcal{O}}(x)$, it defines uniquely the coefficients O^i in $\underline{\mathcal{O}}(x) = O^i\phi_i(x)$ and thus in principle fixes the correlation functions containing arbitrarily many $\underline{\mathcal{O}}(x)$. The ordinary operation $O^i\phi_i(x)$ corresponding to $\underline{\mathcal{O}}(x)$ will be denoted by the same symbol $\underline{\mathcal{O}}(x)$. Define pure-contact operations $\mathcal{O}(x)$ as operations satisfying $\underline{\mathcal{O}}(x) = 0$, i.e.

$$\langle\langle \mathcal{O}(x) = 0 \tag{5.6}$$

Then

$$\begin{aligned}\langle\langle \phi_{i_1}(x_1)\cdots\Theta(y_1)\cdots\mathcal{O}(0)\rangle\rangle = \langle\langle [\phi_{i_1}(x_1), \mathcal{O}(0)]\cdots\Theta(y_1)\cdots\rangle\rangle + \cdots \\ + \langle\langle \phi_{i_1}(x_1)\cdots[\Theta(y_1), \mathcal{O}(0)]\cdots\rangle\rangle + \cdots\end{aligned} \tag{5.7}$$

is a sum of contact terms.

We would like now to construct an operation for a given operator that can be used in computing its correlation functions from $\ln Z$. Since we know how to do this for ordinary operators, it suffices to solve this problem for a pure-contact field. Let $O(x) \in \mathcal{F}$ be a pure-contact field that does not explicitly depend on λ^i, that is $[\partial_i, O(x)] = 0$. Then we can construct a pure-contact operation $\tilde{O}(x)$ by requiring

$$\langle\langle \phi_{i_1}(x_1)\cdots\Theta(y_1)\cdots\tilde{O}(x)\rangle\rangle = \langle \phi_{i_1}(x_1)\cdots\Theta(y_1)\cdots O(x)\rangle_c. \tag{5.8}$$

This essentially fixes $\tilde{O}(x)$ because in physical correlators singularities appear only when some of the insertions coincide. The only ambiguity in $\tilde{O}(x)$ is operations annihilating $\ln Z$. Any choice however suffices for practical purposes. With this definition given an arbitrary operator $A(x) \in \mathcal{F}$ its correlators with the fundamental fields $\phi_{i_k}(x_k)$, $\Theta(y_l)$ can be computed using the ordinary operation $\underline{A}(x) = A^i\phi_i$ and the contact operation

$$\mathcal{A}(x) \equiv [\widetilde{A - \underline{A}}](x) \tag{5.9}$$

according to

$$\langle \phi_{i_1}(x_1)\cdots\Theta(y_1)\cdots A(x)\rangle_c = \langle\langle \phi_{i_1}(x_1)\cdots\Theta(y_1)\cdots[\underline{A}(x) + \mathcal{A}(x)]\rangle\rangle. \tag{5.10}$$

In the above correlation function the contact terms proportional to $\delta(x - x_{i_k})$ are essentially fixed by the action principle (2.4). The extra contributions arising from the explicit dependence of the coefficients A^i on λ^j's are accounted for by commuting the operation $\underline{A}(x)$ to the left. Similarly the contact terms proportional to $\delta(x - y_{i_k})$ are fixed by the change of scale equation (2.3). All contact term contributions proportional to derivatives of delta functions are obtained by commuting the pure-contact operation $\mathcal{A}(x)$ to the left until it annihilates $\langle\langle$.

Consider now the operator $\Theta(x)$. Assuming, as we agreed before, that $\Theta(x) = \beta^i\phi_i \equiv \beta(x)$ in the operator sense means that $\underline{\Theta} = \beta(x)$ and the field $D(x) = \Theta(x) - \beta(x)$ is pure contact. As $\Theta(x)$ does not explicitly depend on λ^i ($\mu\partial/\partial\mu$ and $\partial/\partial\lambda^i$ commute), we can define a pure-contact operation $\mathcal{D}(x)$ in accordance with the general rule (5.9), (5.10). The field $\Theta(x)$ is special in that it is represented by a variational derivative (5.1). This implies that

$$\langle\langle \phi_{i_1}(x_1)\cdots\Theta(y_1)\cdots[\Theta(x) - \beta(x) - \mathcal{D}(x)]\rangle\rangle = 0 \tag{5.11}$$

which can be written more succinctly as a first-order functional differential equation on the generating functional

$$[\Theta(x) - \beta(x) - \mathcal{D}(x)] \ln Z = 0. \tag{5.12}$$

Knowing the pure-contact operation $\mathcal{D}(x)$ the correlation functions of $D(x)$ with any number of $\phi_i(x)$ and $\Theta(x)$ can be calculated as

$$\begin{aligned}
&\langle D(x)\phi_{i_1}(x_1)\ldots\Theta(y_1)\ldots\rangle_c = \langle\langle(\Theta(x)-\beta(x))\phi_{i_1}(x_1)\ldots\Theta(y_1)\cdots\rangle\rangle \\
&\quad = \langle\langle[(\Theta(x)-\beta(x)),\phi_{i_1}(x_1)\ldots\Theta(y_1)\ldots]\rangle\rangle + \langle\langle\phi_{i_1}(x_1)\ldots\Theta(y_1)\ldots\mathcal{D}(x)\rangle\rangle \\
&\quad = \langle\langle[\phi_{i_1}(x_1)\ldots\Theta(y_1)\ldots,(\mathcal{D}(x)+\beta(x)-\Theta(x))]\rangle\rangle \\
&\quad = \langle\langle[\phi_{i_1}(x_1),\mathcal{D}(x)]\ldots\Theta(y_1)\ldots\rangle\rangle + \ldots + \langle\langle\phi_{i_1}(x_1)\ldots[\Theta(y_1),\mathcal{D}(x)]\ldots\rangle\rangle + \ldots \\
&\qquad + \partial_{i_1}\beta^i\delta(x-x_1)\langle\phi_i(x_1)\ldots\Theta(y_1)\ldots\rangle_c + \cdots
\end{aligned} \tag{5.13}$$

where equation (5.11) was used on the second line, $\langle\langle\mathcal{D}(x) = 0$ was used on the third line and

$$[\phi_{i_1}(x_1), \beta(x)] = \delta(x - x_1)\partial_{i_1}\beta^i\phi_i(x_1). \tag{5.14}$$

was used on the last line.

The form of $\mathcal{D}(x)$ is constrained by 2D covariance and locality. In general it can be written as an expansion in derivatives of the sources λ^i and covariant derivatives of the curvature with coefficients being ordinary operations. It is interesting to consider additional restrictions on $\mathcal{D}(x)$ from power counting rules. We will distinguish two such rules which we call a *loose power counting* and a *strict power counting*. In both cases the expansion of $\mathcal{D}(x)$ goes only up to two derivatives in the sources and metric. In the loose power counting rule the coefficients can have a non-trivial operator content. Explicitly in this case we can write

$$\mathcal{D}(x) = \tfrac{1}{2}\mu^2 R_2(x)C(x) + \partial_\mu\lambda^i(x)J_i^\mu(x) + \partial^\mu[W_i(x)\partial_\mu\lambda^i] + \tfrac{1}{2}\partial_\mu\lambda^i\partial^\mu\lambda^j G_{ij}(x) \tag{5.15}$$

where $C(x)$, $W_i(x)$, $G_{ij}(x)$ are ordinary spin-zero fields, and $J_i^\mu(x)$ is an ordinary spin-one field, and where the 2D curvature is given by

$$\mu^2 R_2(x) = -2\partial^\mu\partial_\mu \ln\mu(x).$$

Two comments are in order here. Firstly, note the appearance of vector fields $J_i^\mu(x)$ in the expansion. As we defined operations only for spin-zero fields to accommodate fields and operations of non-trivial spin, we need to introduce new fundamental fields and new sources for those fields. While used to obtain distributional correlation functions involving operators of non-trivial spin, such sources are always set to zero in the end of a computation. The operation $\mathcal{D}(x)$ does contain terms proportional to the tensorial sources and their derivatives. However, our proof avoids using the explicit form of such terms and we will not introduce the tensor field sources explicitly not to clutter the computations. Nevertheless, the operations like $J_i^\mu(x)$, when appear, should be understood in this sense.

Secondly, note that in the power counting scheme used the operators $C(x)$, $W_i(x)$, $G_{ij}(x)$ must have dimension near zero. This means that, using the fixed point language, we allow for slightly irrelevant terms to appear in $\mathcal{D}(x)$. This is a common consideration used for general nonlinear sigma models [5]. The loose power counting thus accommodates perturbative nonlinear sigma models.

If one assumes that the UV behavior is governed by a unitary fixed point, the only dimension zero operator is the identity and the total UV dimension of $\mathcal{D}(x)$ must be strictly 2, then the operators $C(x)$, $W_i(x)$, $G_{ij}(x)$ must be all proportional to the identity operator. We call this restrictions a strict power counting rule. It applies in a vicinity of a unitary fixed point that has a discrete spectrum of conformal dimensions. Under the additional assumption that there are no operators $J_i^\mu(x)$ appearing in $\mathcal{D}(x)$ the case of the strict power counting was investigated in [19].

Finally the case when the only restrictions on $\mathcal{D}(x)$ come from the general covariance and locality can be referred to as Wilsonian. We will prove the general gradient formula (2.13) in the Wilsonian case. The proof is simplified if we impose a loose power counting. We will be discussing in parallel how our steps look in that case.

As a last comment in this section note that due to equation (5.12) the operation $\mathcal{D}(x)$ is subject to Wess–Zumino consistency conditions

$$[\Theta(x) - \beta(x) - \mathcal{D}(x), \Theta(y) - \beta(y) - \mathcal{D}(y)] \ln Z = 0 \tag{5.16}$$

which will be exploited in sections 7.3 and 9.3.

5.2. Summary

Operations $\mathcal{O}(x)$ are local first-order differential operators defined on functionals of the sources $\lambda^i(x)$ and metric. For the fundamental fields $\phi_i(x)$ and the trace of stress–energy tensor $\Theta(x)$ the corresponding operations are the functional derivatives (5.1). We introduced the notation $\langle\langle \mathcal{O}_1(x_1), \ldots, \mathcal{O}(x_n)\rangle\rangle$ for a sequence of operations $\mathcal{O}_i(x_i)$ applied to the generating functional $\ln Z \equiv \rangle\rangle$ with the result restricted to constant sources and metric (the restriction is signified by the symbol $\langle\langle$).

Given an operation $\mathcal{O}(x)$ one can extract a field from it by restricting it to constant sources and metric (5.5). The resulting fields are denoted by $\underline{\mathcal{O}}(x)$ and are called ordinary fields. Such fields have the form $\underline{\mathcal{O}}(x) = O^i\phi_i(x)$. A pure-contact operation is an operation $\mathcal{O}(x)$ for which $\underline{\mathcal{O}}(x) = 0$.

For ordinary fields the distributional correlation functions are completely fixed by those of the fields ϕ_i. More generally a given field $A(x)$ equals a linear combination of fundamental fields: $A(x) = A^i\phi_i(x)$ only up to contact terms. Such contact terms can be stored in a pure-contact operation $\mathcal{A}(x)$ according to (5.10). For the trace of stress–energy tensor $\Theta(x)$ we have $\Theta(x) = \beta^i\phi_i(x) \equiv \beta(x)$ up to contact terms. The corresponding contact terms are stored in a pure-contact operation $\mathcal{D}(x)$. The generating functional satisfies an equation $[\Theta(x) - \beta(x) - \mathcal{D}(x)] \ln Z = 0$ which can be used to compute correlation functions involving the field $D(x) = \Theta(x) - \beta(x)$ according to (5.13). The form of $\mathcal{D}(x)$ is constrained by locality and general covariance. It can be further constrained by a power counting principle. We distinguish a strict power counting, which applies to a vicinity of a unitary fixed point with a discrete spectrum of conformal dimensions, and a loose power counting that is suitable for describing renormalizable nonlinear sigma models. For the loose power counting case $\mathcal{D}(x)$ can be explicitly written as in formula (5.15).

6. The Callan–Symanzik equations

In the operations formalism the Callan–Symanzik equations for correlators involving fields $\phi_i(x)$ and $\Theta(y)$ can be obtained by integrating equation (5.13) over x:

$$\begin{aligned}\left(\mu\frac{\partial}{\partial\mu} - \beta^i\frac{\partial}{\partial\lambda^i}\right)\langle\phi_{i_1}(x_1)\ldots\Theta(y_1)\ldots\rangle_c &= \int \mathrm{d}^2x\, \langle\langle D(x)\,\phi_{i_1}(x_1)\ldots\Theta(y_1)\ldots\rangle\rangle \\ &= \partial_{i_1}\beta^i\langle\phi_i(x_1)\ldots\Theta(y_1)\ldots\rangle_c + \int \mathrm{d}^2x\,\langle\langle[\phi_{i_1}(x_1), \mathcal{D}(x)]\ldots\Theta(y_1)\rangle\rangle + \ldots \\ &\quad + \int \mathrm{d}^2x\,\langle\langle\phi_{i_1}(x_1)\ldots[\Theta(y_1), \mathcal{D}(x)]\ldots\rangle\rangle + \ldots.\end{aligned} \tag{6.1}$$

It is convenient to define the following operations:

$$\mathcal{D}\phi_i(x) = \int d^2y[\phi_i(x), \mathcal{D}(y)],$$

$$\mathcal{D}\Theta(x) = \int d^2y[\Theta(x), \mathcal{D}(y)]. \tag{6.2}$$

In view of (6.1) the operations $\mathcal{D}\phi_i(x)$ and $\mathcal{D}\Theta(x)$ can be interpreted as extra contributions to the Callan–Symanzik equations.

We further note that

$$\int d^2x \langle\langle \mathcal{D}\Theta(x) = 0, \qquad \int d^2x \langle\langle \mathcal{D}\phi_i(x) = 0. \tag{6.3}$$

This follows from the fact that $\int d^2x\, \Theta(x) = \mu\partial/\partial\mu$, $\int d^2y\, \phi_i(x) = \partial/\partial\lambda^i$ and every term in $\mathcal{D}(y)$ is proportional to derivatives of $\lambda^i(y)$ and $\mu(y)$. Equations (6.3) imply that there must be ordinary spin-one fields (and ordinary operations respectively) $J^\mu(x)$ and $J_i^\mu(x)$ such that

$$\underline{\mathcal{D}\Theta}(x) = -\partial_\mu J^\mu(x), \tag{6.4}$$

$$\underline{\mathcal{D}\phi_i}(x) = -\partial_\mu J_i^\mu(x). \tag{6.5}$$

If we impose a loose power counting, so that $\mathcal{D}(x)$ is given by equation (5.15), then

$$\mathcal{D}\Theta(x) = -\partial^\mu\partial_\mu C(x) \tag{6.6}$$

$$\mathcal{D}\phi_i(x) = -\partial_\mu\left[J_i^\mu(x) + \partial^\mu\lambda^j G_{ij}(x)\right] + \partial_\mu\lambda^j\partial_i J_j^\mu(x) + \tfrac{1}{2}\partial_\mu\lambda^j\partial^\mu\lambda^k\partial_i G_{jk}(x) \tag{6.7}$$

so

$$J^\mu(x) = \partial^\mu C(x) \tag{6.8}$$

and $J_i^\mu(x)$ defined in (6.5) in general (without any power counting assumptions) coincides with the coefficient in the expansion of $\mathcal{D}(x)$ based on a loose power counting, equation (5.15). In general (without any power counting restrictions) since all terms in $\mathcal{D}(x)$ are proportional to derivatives of the sources and/or to derivatives of $\mu(x)$, there exists a scalar operator $C(x)$ such that $J^\mu(x) = \partial^\mu C(x)$.

The Callan–Symanzik equations (6.1) for the correlation functions at non-coincident points (neglecting contact terms) can now be written as

$$\begin{aligned}\mu\frac{\partial}{\partial\mu}\langle\phi_{i_1}(x_1)\ldots\Theta(y_1)\ldots\rangle_c = \beta^i\frac{\partial}{\partial\lambda^i}\langle\phi_{i_1}(x_1)\cdots\Theta(y_1)\ldots\rangle_c + \langle\Gamma\phi_{i_1}(x_1)\ldots\Theta(y_1)\ldots\rangle_c\\ +\ldots+\langle\phi_{i_1}(x_1)\ldots[-\partial_\mu J^\mu(y_1)]\ldots\rangle_c+\ldots\end{aligned} \tag{6.9}$$

where

$$\Gamma\phi_{i_1}(x_1) = \partial_{i_1}\beta^i\phi_i(x_1) - \partial_\mu J_{i_1}^\mu(x_1). \tag{6.10}$$

The terms involving the beta functions can be put into the Lie derivative $\mathcal{L}_\beta$ so that equation (6.9) takes a more succinct form

$$\begin{aligned}&\left[\mu\frac{\partial}{\partial\mu} - \mathcal{L}_\beta\right]\langle\phi_{i_1}(x_1)\cdots\Theta(y_1)\ldots\rangle_c\\ &\quad = \big\langle\big[-\partial_\mu J_{i_1}^\mu(x_1)\big]\ldots\Theta(y_1)\ldots\big\rangle_c+\ldots+\langle\phi_{i_1}(x_1)\ldots[-\partial_\mu J^\mu(y_1)]\ldots\rangle_c+\ldots.\end{aligned} \tag{6.11}$$

7. The proof continued

We now come back to the proof of the gradient formula which we left at the end of section 4. We express the remainder term r_i^L of equation (4.12) in the source-operation formalism. The three-point functions occurring in equation 4.12 can be written as

$$\begin{aligned}\langle\phi_i(y)D(x)\Theta(0)\rangle_c &= \langle\langle\phi_i(y)\Theta(0)[\Theta(x)-\beta(x)]\rangle\rangle + \partial_i\beta^j\delta^2(y-x)\langle\langle\Theta(0)\phi_j(x)\rangle\rangle\\ &= \langle\langle\phi_i(y)\Theta(0)\mathcal{D}(x)\rangle\rangle + \partial_i\beta^j\delta^2(y-x)\langle\langle\Theta(0)\phi_j(x)\rangle\rangle, \end{aligned} \tag{7.1}$$

$$\langle D(y)\phi_i(x)\Theta(0)\rangle_c = \langle\langle\phi_i(x)\Theta(0)\mathcal{D}(y)\rangle\rangle + \partial_i\beta^j\delta^2(x-y)\langle\langle\Theta(0)\phi_j(y)\rangle\rangle, \tag{7.2}$$

so

$$\langle\phi_i(y)D(x)\Theta(0) - D(y)\phi_i(x)\Theta(0)\rangle_c = \langle\langle\phi_i(y)\Theta(0)\mathcal{D}(x) - \phi_i(x)\Theta(0)\mathcal{D}(y)\rangle\rangle. \tag{7.3}$$

Substituting the last relation into equation (4.12) and using $\langle\langle\,\mathcal{D}(x) = 0$ gives

$$\begin{aligned} r_i^L &= \int_{|y|<L} \mathrm{d}^2y \int \mathrm{d}^2x G_\Lambda(x)\langle\langle\phi_i(y)\Theta(0)\mathcal{D}(x) - \phi_i(x)\Theta(0)\mathcal{D}(y)\rangle\rangle\\ &= \int_{|y|<L} \mathrm{d}^2y \int \mathrm{d}^2x G_\Lambda(x)\langle\langle\phi_i(y)[\Theta(0),\mathcal{D}(x)] + [\phi_i(y),\mathcal{D}(x)]\Theta(0)\rangle\rangle\\ &\quad - \int_{|y|<L} \mathrm{d}^2y \int \mathrm{d}^2x G_\Lambda(x)\langle\langle\phi_i(x)[\Theta(0),\mathcal{D}(y)] + [\phi_i(x),\mathcal{D}(y)]\Theta(0)\rangle\rangle. \end{aligned}\tag{7.4}$$

Note that $\mathcal{D}(x)$ is a pure-contact operation, and $|x| \leqslant \Lambda^{-1} \ll L$, so that

$$\int_{|y|<L} \mathrm{d}^2y\ [\Theta(0),\mathcal{D}(y)] = \int \mathrm{d}^2y\ [\Theta(0),\mathcal{D}(y)] = \mathcal{D}\Theta(0) \tag{7.5}$$

$$\int_{|y|<L} \mathrm{d}^2y\ [\phi_i(x)\mathcal{D}(y)] = \int \mathrm{d}^2y\ [\phi_i(x)\mathcal{D}(y)] = \mathcal{D}\phi_i(0) \tag{7.6}$$

$$\int_{|y|<L} \mathrm{d}^2y\ \langle\langle[\phi_i(y),\mathcal{D}(x)] = \int \mathrm{d}^2y\ \langle\langle[\phi_i(y),\mathcal{D}(x)] = [\partial_i,\mathcal{D}(x)] = 0. \tag{7.7}$$

Using these relations in (7.4) we obtain

$$\begin{aligned} r_i^L = &-\int_{|y|<L} \mathrm{d}^2y 12\pi\langle\langle\phi_i(y)C_2(0)\rangle\rangle - \int \mathrm{d}^2x G_\Lambda(x)\langle\langle\phi_i(x)\mathcal{D}\Theta(0)\rangle\rangle\\ &-\int \mathrm{d}^2x G_\Lambda(x)\langle\langle\mathcal{D}\phi_i(x)\Theta(0)\rangle\rangle \end{aligned}\tag{7.8}$$

where we have defined an operation

$$C_2(y) = -\int \mathrm{d}^2x \frac{1}{4}x^2[\Theta(y),\mathcal{D}(x)]. \tag{7.9}$$

If a loose power counting is imposed, $\mathcal{D}(x)$ is given by equation (5.15), and we have

$$C_2(y) = -\int \mathrm{d}^2x \frac{1}{4}x^2[-\partial^\mu\partial_\mu\delta^2(x-y)]C(x) = C(y). \tag{7.10}$$

Thus, with a loose power counting,

$$\mathcal{D}\Theta(x) = -\partial^\mu\partial_\mu C_2(x). \tag{7.11}$$

We separate r_i^L into two parts:

$$r_i^L = r_{i,1}^L + r_{i,2}^L \tag{7.12}$$

$$r_{i,1}^L = -\int d^2x\, G_\Lambda(x)\langle\langle \mathcal{D}\phi_i(x)\Theta(0)\rangle\rangle \tag{7.13}$$

$$r_{i,2}^L = -\int_{|y|<L} d^2y\, 12\pi\,\langle\langle \phi_i(y) C_2(0)\rangle\rangle - \int d^2x\, G_\Lambda(x)\langle\langle \phi_i(x)\mathcal{D}\Theta(0)\rangle\rangle \tag{7.14}$$

and then investigate each in turn.

7.1. The IR condition and the sum rule

We investigate $r_{i,1}$ first. Our goal is to show that under certain assumptions this quantity is proportional to the beta functions.

We have

$$\langle\langle \mathcal{D}\phi_i(x) = \langle\langle \underline{\mathcal{D}\phi_i}(x) = \langle\langle\big[-\partial_\mu J_i^\mu(x)\big] \tag{7.15}$$

so

$$\langle\langle \mathcal{D}\phi_i(x)\Theta(0)\rangle\rangle = -\langle\langle \partial_\mu J_i^\mu(x)\Theta(0)\rangle\rangle = -\big\langle \partial_\mu J_i^\mu(x)\Theta(0)\big\rangle_c. \tag{7.16}$$

Substituting this expression into equation (7.13) we get

$$r_{i,1}^L = \int d^2x\, G_\Lambda(x)\big\langle \partial_\mu J_i^\mu(x)\Theta(0)\big\rangle_c. \tag{7.17}$$

Now we use the technique similar to the one we used in the proof of Zamolodchikov's formula (see section 3). It is straightforward to check that the Ward identity for $T_{\mu\nu}(x)$ implies

$$x^2\langle\partial_\mu J_i^\mu(x)\Theta(0)\rangle_c = \partial_\mu\big[x^2\big\langle J_i^\mu(x)\Theta(0)\big\rangle_c - 2x_\alpha x^\beta\big\langle J_i^\alpha(x)T_\beta^\mu(0)\big\rangle_c + x^2\big\langle J_i^\alpha(x)T_\alpha^\mu(0)\big\rangle_c\big] \tag{7.18}$$

which allows us to perform the integral in equation (7.17), obtaining

$$r_{i,1}^L = 6\pi^2 x_\mu\big[x^2\big\langle J_i^\mu(x)\Theta(0)\big\rangle_c - 2x_\alpha x^\beta\big\langle J_i^\alpha(x)T_\beta^\mu(0)\big\rangle_c + x^2\big\langle J_i^\alpha(x)T_\alpha^\mu(0)\big\rangle_c\big]_{/\Lambda|x|=1}. \tag{7.19}$$

What we want however is an expression proportional to β^i. Recall that

$$G_\Lambda(x) = 3\pi x^2\theta(1-\Lambda|x|) \tag{7.20}$$

so that

$$G_0(x) = 3\pi x^2 \tag{7.21}$$

and

$$G_0(x) - G_\Lambda(x) = 3\pi x^2\theta(\Lambda|x|-1). \tag{7.22}$$

We write

$$\begin{aligned} r_{i,1}^L &= E_2 + \int_{|x|\leqslant L} d^2x[G_\Lambda(x) - G_0(x)]\big\langle \partial_\mu J_i^\mu(x)\Theta(0)\big\rangle_c \\ &= E_2 + \int_{|x|\leqslant L} d^2x[G_\Lambda(x) - G_0(x)]\big\langle \partial_\mu J_i^\mu(x)\phi_j(0)\big\rangle_c\beta^j \end{aligned} \tag{7.23}$$

with

$$E_2 = \int_{|x|\leqslant L} \mathrm{d}^2x G_0(x)\big\langle \partial_\mu J_i^\mu(x)\Theta(0)\big\rangle_c \tag{7.24}$$

$$= 6\pi^2 x_\mu\big[x^2\big\langle J_i^\mu(x)\Theta(0)\big\rangle_c - 2x_\alpha x^\beta\big\langle J_i^\alpha(x)T^\mu_\beta(0)\big\rangle_c + x^2\big\langle J_i^\alpha(x)T^\mu_\alpha(0)\big\rangle_c\big]_{/|x|=L}. \tag{7.25}$$

We are allowed to replace $\Theta(0)$ with $\beta^j\phi_j(0)$ to obtain equation (7.23) because $G_\Lambda(x)-G_0(x)$ vanishes for $\Lambda|x| \leqslant 1$, so contact terms in the two-point function make no difference.

The IR error term E_2 will vanish in the limit $L \to \infty$ if the two-point functions $\langle J_i^\mu(x)\, T_{\alpha\beta}(0)\rangle_c$ go to zero at large x faster than $|x|^{-3}$:

$$\lim_{|x|\to\infty} |x|^3\big\langle J_i^\mu(x)\, T_{\alpha\beta}(0)\big\rangle_c = 0. \tag{7.26}$$

A violation of this IR decay condition would mean that the long-distance limit of the quantum field theory exhibits spontaneously broken global conformal symmetry. Our main IR assumption is that such a spontaneous breaking does not take place and equation (7.26) is satisfied.

The condition $\lim_{L\to\infty} E_2 = 0$ is equivalent to the sum rule

$$\int \mathrm{d}^2x x^2\big\langle \partial_\mu J_i^\mu(x)\Theta(0)\big\rangle_c = 0. \tag{7.27}$$

Such a sum rule holds for any spin-one field, given our infrared assumption.

7.2. *The term $r^L_{i,2}$*

Similar to (7.23) we want to write $r^L_{i,2}$ as an integral over $\Lambda|x| > 1$ of an expression proportional to β^j. Equation (7.11), which one obtains when the loose power counting is imposed, motivates the following manipulation of equation (7.14). Write the first term, using equation (7.21) for $G_0(y)$,

$$-\int_{|y|<L} \mathrm{d}^2y 12\pi\langle\langle\phi_i(y)C_2(0)\rangle\rangle = -\int_{|y|<L} \mathrm{d}^2y[\partial_\mu\partial^\mu G_0(y)]\langle\langle\phi_i(y)C_2(0)\rangle\rangle \tag{7.28}$$

and then integrate by parts. Equation (7.14) becomes

$$r^L_{i,2} = E_3 - \int_{|y|<L} \mathrm{d}^2y G_0(y)\langle\langle\phi_i(y)\partial_\mu\partial^\mu C_2(0)\rangle\rangle - \int \mathrm{d}^2x G_\Lambda(x)\langle\langle\phi_i(x)\mathcal{D}\Theta(0)\rangle\rangle \tag{7.29}$$

where E_3 is an infrared error

$$E_3 = -\int_{|x|<L} \mathrm{d}^2x\partial_\mu[\partial^\mu G_0(x)\langle\langle\phi_i(0)C_2(x)\rangle\rangle - G_0(x)\partial^\mu\langle\langle\phi_i(0)C_2(x)\rangle\rangle]. \tag{7.30}$$

We further rewrite equation (7.29) as

$$r^L_{i,2} = E_3 + E_4 + \int_{|x|<L} \mathrm{d}^2x\,[G_0(x) - G_\Lambda(x)]\,\langle\langle\phi_i(x)\mathcal{D}\Theta(0)\rangle\rangle \tag{7.31}$$

where

$$E_4 = -\int_{|x|<L} \mathrm{d}^2x G_0(x)\langle\langle\phi_i(x)[\partial_\mu\partial^\mu C_2(0) + \mathcal{D}\Theta(0)]\rangle\rangle. \tag{7.32}$$

The term E_4 is identically zero if we assume a loose power counting by equation (7.11). We will show in section 7.4 that in general E_4 vanishes as $L \to \infty$.

In equation (7.31), the integration variable x is bounded away from 0, so we can substitute

$$\langle\langle\phi_i(x)\mathcal{D}\Theta(0)\rangle\rangle = \langle\langle\mathcal{D}\Theta(0)\phi_i(x)\rangle\rangle = \langle -\partial_\mu J^\mu(0)\phi_i(x)\rangle_c \tag{7.33}$$

giving

$$r^L_{i,2} = E_3 + E_4 + \int_{|x|<L} \mathrm{d}^2x\,[G_\Lambda(x) - G_0(x)]\,\langle\phi_i(x)\partial_\mu J^\mu(0)\rangle_c. \tag{7.34}$$

Finally, we will now show that

$$\partial_\mu J^\mu(0) = \beta^j\partial_\mu J^\mu_j(0) \tag{7.35}$$

so that $r^L_{i,2}$ also becomes proportional to β^j, up to IR errors,

$$r^L_{i,2} = E_3 + E_4 + \int_{|x|<L} \mathrm{d}^2x\,[G_\Lambda(x) - G_0(x)]\,\langle\phi_i(x)\partial_\mu J^\mu_j(0)\rangle_c\beta^j. \tag{7.36}$$

7.3. The identity $\partial_\mu J^\mu(x) = \beta^j\partial_\mu J^\mu_j(x)$

We want to show that the ordinary field

$$K(x) = \beta^j\partial_\mu J^\mu_j(x) - \partial_\mu J^\mu(x) \tag{7.37}$$

is zero, which is to say that all its non-coincident correlation functions vanish:

$$\langle K(x)\phi_{i_1}(x_1)\ldots\rangle_c = 0 \qquad x \neq x_1,\ldots. \tag{7.38}$$

In the source/operation formalism, this means that

$$\langle\langle K(x)\phi_{i_1}(x_1)\ldots\rangle\rangle = 0, \qquad x \neq x_1,\ldots. \tag{7.39}$$

To show this we first argue that (7.38) is equivalent to showing that

$$[\mathcal{D}, D(x)]\rangle\rangle = \mathcal{K}_1(x)\rangle\rangle \tag{7.40}$$

for some pure-contact operation $\mathcal{K}_1(x)$. We then demonstrate that (7.40) is a consequence of the Wess–Zumino consistency conditions on $\mathcal{D}(x)$.

It follows from (6.2) that

$$\begin{aligned}\langle\langle K(x) &= \langle\langle\big[-\partial_\mu J^\mu(x) + \beta^j\partial_\mu J^\mu_j(x)\big]\\ &= \langle\langle[\mathcal{D}\Theta(x) - \beta^j\mathcal{D}\phi_j(x)]\\ &= \langle\langle[\mathcal{D}, \Theta(x) - \beta^j\phi_j(x)]\\ &= \langle\langle[\mathcal{D}, D(x)]\end{aligned} \tag{7.41}$$

where $D(x) = \Theta(x) - \beta^j\phi_j(x)$ is acting here as an operation. This last calculation implicitly uses the obvious identity

$$\langle\langle\beta^j(\lambda(x)) = \beta^j(\lambda)\langle\langle \tag{7.42}$$

and its direct implication

$$\langle\langle[\mathcal{D}, \beta^j(\lambda(x))] = -\langle\langle\beta^j(\lambda(x))\mathcal{D} = -\beta^j(\lambda)\langle\langle\mathcal{D} = 0. \tag{7.43}$$

Now we have

$$\langle\langle K(x)\phi_{i_1}(x_1)\cdots\rangle\rangle = \langle\langle[\mathcal{D}, D(x)]\phi_{i_1}(x_1)\cdots\rangle\rangle. \tag{7.44}$$

The operation $[\mathcal{D},\ D(x)]$ commutes with all $\phi_{i_r}(x_r)$ because $x \neq x_r$, so

$$\langle\langle K(x)\phi_{i_1}(x_1)\cdots\rangle\rangle = \langle\langle \phi_{i_1}(x_1)\cdots[\mathcal{D}, D(x)]\rangle\rangle. \tag{7.45}$$

We now need to show that

$$\langle\langle \phi_{i_1}(x_1)\cdots[\mathcal{D}, D(x)]\rangle\rangle = 0, \qquad x \neq x_1, \ldots, \tag{7.46}$$

which by (5.7) is equivalent to (7.40).

Equation (7.40) follows from the Wess–Zumino consistency conditions. Recall that we have an equation

$$0 = [D(x) - \mathcal{D}(x)]\rangle\rangle. \tag{7.47}$$

The Wess–Zumino consistency conditions are

$$[D(x) - \mathcal{D}(x), D(y) - \mathcal{D}(y)]\rangle\rangle = 0. \tag{7.48}$$

It follows from

$$[\Theta(x), \Theta(y)] = 0, \qquad [\Theta(x), \beta(y)] = 0, \qquad [\beta(x), \beta(y)] = 0 \tag{7.49}$$

that

$$[D(x), D(y)] = 0 \tag{7.50}$$

and therefore (7.47) is equivalent to

$$[\mathcal{D}(y), D(x)]\rangle\rangle = -([D(y), \mathcal{D}(x)] + [\mathcal{D}(x), \mathcal{D}(y)])\rangle\rangle. \tag{7.51}$$

The operation $[\mathcal{D}(x),\ \mathcal{D}(y)]$ is evidently pure contact. It also follows from (2.3) and (7.43) that

$$\left[\int \mathrm{d}^2 y D(y), \mathcal{D}(x)\right] \tag{7.52}$$

is a pure-contact operation. Thus, integrating equation (7.51) with respect to y gives

$$[\mathcal{D}, D(x)]\rangle\rangle = \mathcal{K}_1(x)\rangle\rangle \tag{7.53}$$

where

$$\mathcal{K}_1(x) = -\left[\int \mathrm{d}^2 y D(y), \mathcal{D}(x)\right] - [\mathcal{D}(x), \mathcal{D}] \tag{7.54}$$

is pure contact. This completes the proof that at all non-coincident correlation functions of $\beta^j \partial_\mu J^\mu_j(x) - \partial_\mu J^\mu(x)$ are identically zero. Therefore,[6]

$$\partial_\mu J^\mu(x) = \beta^j \partial_\mu J^\mu_j(x). \tag{7.55}$$

7.4. E_4 is an IR error term

We owe a proof that the term E_4 given by

$$E_4 = -\int_{|x|<L} \mathrm{d}^2 x\, G_0(x)\langle\langle \phi_i(x)[\partial_\mu \partial^\mu C_2(0) + \mathcal{D}\Theta(0)]\rangle\rangle \tag{7.56}$$

is an infrared error, that is it vanishes as $L \to \infty$. The argument is a bit tedious, so the reader might want to skip this section at the first reading.

[6] It is worth noting that relation (7.55) is a generalization of the Curci–Paffuti relation [16] known for nonlinear sigma models. By methods similar to those employed in this section one can actually prove a stronger relation: $J^\mu(x) = \beta^j J^\mu_j(x)$. We do not need this stronger relation in the proof of the gradient formula.

We have noted that in general (without the assumption of a loose power counting) we have

$$[\Theta(0), \mathcal{D}(y)] = -\partial_\mu\partial^\mu\delta(y)\, C_2(y) + \partial_\mu\partial_\nu\partial_\gamma\delta(y)\, C_3^{\mu\nu\gamma}(y) + \cdots \tag{7.57}$$

where the omitted terms contain derivatives of delta functions of order 4 and higher. For our purposes this expansion can be written more compactly as

$$[\Theta(0), \mathcal{D}(y)] = -\partial_\mu\partial^\mu\delta(y)\, C_2(y) + \partial_\mu\partial_\nu\partial_\gamma\delta(y)\, \tilde{C}_3^{\mu\nu\gamma}(y) \tag{7.58}$$

where $\tilde{C}_3^{\mu\nu\gamma}(y)$ is some tensor operation. Formula (7.58) implies

$$\partial_\mu\partial^\mu C_2(0) + \mathcal{D}\Theta(0) = \partial_\mu\partial_\nu\partial_\gamma\tilde{C}_3^{\mu\nu\gamma}(0) \tag{7.59}$$

and therefore

$$\langle\langle\phi_i(x)[\partial_\mu\partial^\mu C_2(0) + \mathcal{D}\Theta(0)]\rangle\rangle = \left\langle\partial_\mu\partial_\nu\partial_\gamma\underline{\tilde{C}_3}^{\mu\nu\gamma}(0)\phi_i(x)\right\rangle + \left\langle\!\left\langle\left[\phi_i(x), \partial_\mu\partial_\nu\partial_\gamma\tilde{C}_3^{\mu\nu\gamma}(0)\right]\right\rangle\!\right\rangle. \tag{7.60}$$

The second term on the right-hand side of (7.60) vanishes because it is proportional to a one-point function of a total derivative operator. Thus, we obtain

$$E_4 = -3\pi\int_{|x|<L} \mathrm{d}^2x\, x^2\left\langle\phi_i(x)\partial_\mu\partial_\nu\partial_\gamma\underline{\tilde{C}_3}^{\mu\nu\gamma}(0)\right\rangle \tag{7.61}$$

which exhibits that E_4 is a linear combination of two-point functions at separation L. Assuming that $\langle\phi_i(L)\underline{\tilde{C}_3}^{\mu\nu\gamma}(0)\rangle$ is integrable at infinity (which is consistent with $\langle\underline{\tilde{C}_3}^{\mu\nu\gamma}(0)\rangle = 0$ being independent of λ_i) all combinations of two-point functions entering E_4 go to zero as $L \to \infty$.

8. Conclusion of the proof

Combining our results for $r_{i,1}^L$ and $r_{i,2}^L$, equations (7.23) and (7.36), and substituting into equation (7.12), we get

$$r_i^L = E_2 + E_3 + E_4 + \left(\Delta g_{ij}^L\right)\beta^j \tag{8.1}$$

with

$$\Delta g_{ij}^L = \int_{|x|<L} \mathrm{d}^2x[G_\Lambda(x) - G_0(x)]\left\langle\phi_i(x)\partial_\mu J_j^\mu(0) + \phi_j(x)\partial_\mu J_i^\mu(0)\right\rangle_c. \tag{8.2}$$

The metric correction Δg_{ij}^L can be also written without any direct reference to currents $J_i^\mu(x)$ using the Callan–Symanzik equations (6.11)

$$\Delta g_{ij}^L = \int_{|x|<L} \mathrm{d}^2x\, [G_\Lambda(x) - G_0(x)]\left(\mathcal{L}_\beta - \mu\frac{\partial}{\partial\mu}\right)\langle\phi_i(x)\phi_j\rangle \tag{8.3}$$

which, using (2.3), (2.4), can be written in terms of integrated three-point functions of fundamental operators up to IR error terms.

Equation (4.11) becomes, finally, the IR-regulated gradient formula

$$\partial_i^L c + \left(g_{ij} + \Delta g_{ij}^L + b_{ij}^L\right)\beta^j + E(L) = 0 \tag{8.4}$$

with total error

$$E(L) = E_1 + E_2 + E_3 + E_4. \tag{8.5}$$

The L-dependent constituents of the formula are

$$\partial_i^L c = -\int_{|y|<L} \mathrm{d}^2 y \int \mathrm{d}^2 x\ G_\Lambda(x)\langle\phi_i(y)\Theta(x)\Theta(0)\rangle_c, \tag{8.6}$$

$$b_{ij}^L = \int_{|y|<L} \mathrm{d}^2 y \int \mathrm{d}^2 x G_\Lambda(x)\langle\phi_i(y)\phi_j(x)\Theta(0) - \phi_j(y)\phi_i(x)\Theta(0)\rangle_c, \tag{8.7}$$

$$E_1 = 2\pi y^\mu y^\nu \int \mathrm{d}^2 x\ G_\Lambda(x)\langle T_{\mu\nu}(y)\phi_i(x)\Theta(0)\rangle_{c/|y|=L}, \tag{8.8}$$

$$E_2 = 6\pi^2 x_\mu\left[x^2\left\langle J_i^\mu(x)\Theta(0)\right\rangle_c - 2x_\alpha x^\beta\left\langle J_i^\alpha(x)T_\beta^\mu(0)\right\rangle_c + x^2\left\langle J_i^\alpha(x)T_\alpha^\mu(0)\right\rangle_c\right]_{/|x|=L}, \tag{8.9}$$

$$E_3 = -\int_{|x|<L} \mathrm{d}^2 x \partial_\mu[\partial^\mu G_0(x)\langle\langle\phi_i(0)C_2(x)\rangle\rangle - G_0(x)\partial^\mu\langle\langle\phi_i(0)C_2(x)\rangle\rangle], \tag{8.10}$$

$$E_4 = -3\pi\int_{|x|<L} \mathrm{d}^2 x x^2\left\langle\phi_i(x)\partial_\mu\partial_\nu\partial_\gamma \underline{\tilde{C}_3}^{\mu\nu\gamma}(0)\right\rangle \tag{8.11}$$

and Δg_{ij}^L is given in (8.2) (see equations (4.8), (4.10), (4.6), (7.25), (7.30) and (7.61)).

Now that the infrared regulated formula (8.4) is derived we can study its $L \to \infty$ limit. Let us recapitulate our assumptions on the infrared behavior. Firstly, we assume that the action principle holds at least for one- and two-point functions so that the one- and two-point functions are at least once differentiable. Secondly, the infrared behavior of the stress–energy tensor correlators should satisfy (7.26). The first assumption means that two-, three- and four-point functions involving $\phi_i(x)$ or $T_{\mu\nu}(x)$ decay faster than x^2 when $|x| \to \infty$. This together with formula (7.26) implies that

$$\begin{aligned}
&\lim_{L\to\infty} E(L) = 0,\\
&\lim_{L\to\infty} \partial_i^L c = \partial_i c,\\
&\lim_{L\to\infty} b_{ij}^L = b_{ij},
\end{aligned} \tag{8.12}$$

where b_{ij} is given by Osborn's formula[7] (2.12). Note that in showing (8.12) formula (7.26) is needed only to argue that E_2 vanishes at infinity while the first infrared assumption alone suffices to show all other limits.

Note that although the same set of assumptions implies

$$\lim_{L\to\infty} \Delta g_{ij}^L \beta^j < \infty, \tag{8.13}$$

there is no guarantee that the $L \to \infty$ limit of Δg_{ij}^L is finite. However, infrared divergences, if present in Δg_{ij}^L, are orthogonal to the beta function. Therefore, they can be subtracted to obtain a finite quantity Δg_{ij} so that the following gradient formula holds:

$$\partial_i c = -(g_{ij} + \Delta g_{ij} + b_{ij})\beta^j \tag{8.14}$$

[7] The two-form b_{ij} is exact provided w_j defined in (2.12) is differentiable. If one relaxes the differentiability assumptions, there is room for the limit $b_{ij} = \lim_{L\to\infty} b_{ij}^L$ to exist without w_j being differentiable, in which case b_{ij} would be closed but not exact. The failure of differentiability of w_j could come from some non-perturbative effects.

where

$$\Delta g_{ij} = \lim_{L\to\infty} \left[\Delta g_{ij}^L - \text{ subtractions}\right]. \tag{8.15}$$

This completes the derivation of the general gradient formula.

9. Discussion

9.1. Contact term ambiguities and scale dependence

As the proof of the gradient formula uses distributional correlation functions which have contact term ambiguities, one should ask if the formula itself is free from such ambiguities. The contact term ambiguities arise from the choice of a renormalization scheme and are generated by adding to the generating functional finite local counterterms of the form

$$\ln Z[\lambda, g_{ij}] \mapsto \ln Z[\lambda, g_{ij}] + \int \mathrm{d}^2x \left[f(\lambda)\mu^2 R_2(x) + \frac{1}{2}c_{ij}(\lambda)\partial_\mu\lambda\partial^\mu\lambda(x) + \cdots\right], \tag{9.1}$$

where $f(\lambda)$ and $c_{ij}(\lambda)$ are arbitrary functions[8] (scalar and tensor respectively) and the omitted terms contain higher order derivatives of the metric and sources. Redefinition (9.1) shifts the terms in the renormalization operation $\mathcal{D}(x)$. The low-order terms shift as

$$C(x) \mapsto C(x) + \beta^i\partial_i f(x), \tag{9.2}$$

$$W_i(x) \mapsto W_i(x) - \partial_i f(x) - c_{ij}\beta^j(x), \tag{9.3}$$

$$G_{ij}(x) \mapsto G_{ij}(x) - \mathcal{L}_\beta c_{ij}(x) \tag{9.4}$$

with all shifts proportional to the identity operator.

The c-function and the metric tensors g_{ij}, Δg_{ij} can each be written in a form involving two-point correlators at non-zero separation only (see formulas (2.10), (2.11), (8.2)). Thus, these quantities are independent of the contact term ambiguities. The one-form w_i defined in (48) changes under (9.1) as

$$w_i \mapsto w_i - \partial_i f \tag{9.5}$$

and the antisymmetric form b_{ij} thus does not change. Since redefinition (9.1) is the most general one[9], the two-form b_{ij} is also independent of the contact term ambiguities.

Another property that we would like to check is whether the quantities we defined depend on the scales μ and Λ only via their ratio μ/Λ. For the c-function (2.10), the metric (2.11) and the antisymmetric form (2.12) this immediately follows from the scaling properties (3.2). As for the metric correction Δg_{ij} it may happen that the infrared regulated quantity Δg_{ij}^L contains a logarithmic divergence $\sim \ln L$ whose subtraction requires introducing a new scale. If this happens, the subtracted correction will not depend on μ and Λ via the ratio μ/Λ only. The physical significance of this is unclear to us.

9.2. The infrared condition: an example

Here we discuss a simple example that demonstrates the necessity of the infrared condition (7.26) for a gradient formula to hold. Consider a free compact boson X defined on a two-dimensional curved surface with metric $g_{\mu\nu}$ by the action functional

$$S[R, g_{\mu\nu}] = \frac{1}{8\pi}\int \mathrm{d}^2x\, (\lambda\sqrt{g}g^{\mu\nu}\partial_\mu X\partial_\nu X + QX\sqrt{g}R_2), \tag{9.6}$$

[8] We assume that these functions are at least once differentiable.

[9] The higher order terms omitted in (9.1) do not contribute to the change of w_i.

where λ is the coupling constant corresponding to the radius of compactification squared, R_2 is the curvature of $g_{\mu\nu}$ and Q is a parameter. Promoting λ to a local source $\lambda(x)$ we can define a generating functional

$$\ln Z[\lambda(x), g_{\mu\nu}(x)] = \int [dX]\, e^{-S[\lambda(x), g_{\mu\nu}(x)]}. \tag{9.7}$$

For the zero mode integral to be well defined we assume that the theory is defined only on a surface with the topology of a plane so that

$$\int d^2x\, \sqrt{g} R_2 = 0 \tag{9.8}$$

and the zero mode integral in (9.7) only yields an overall numerical factor. Note that Q cannot be considered as a coupling constant as it does not stand at a local operator. The functional integral is Gaussian so the anomaly can be readily computed (e.g. using the heat kernel method) with the result

$$D(x) = \Theta(x) = \tfrac{1}{2}C(\lambda)\sqrt{g}R_2(x) + J^{\mu}_{\lambda}(x)\partial_\mu\lambda + \tfrac{1}{2}g_{\lambda\lambda}\partial_\mu\lambda\partial^\mu\lambda + \partial_\mu(w_\lambda\partial^\mu\lambda), \tag{9.9}$$

where

$$C(\lambda) = \frac{1}{12\pi} + \frac{Q^2}{4\pi\lambda} \tag{9.10}$$

$$J^{\mu}_{\lambda}(x) = -\frac{Q}{4\pi\lambda}\partial^\mu X(x) \tag{9.11}$$

$$g_{\lambda\lambda} = \frac{1}{64\pi\lambda^2}. \tag{9.12}$$

The value of w_λ is essentially scheme dependent. It can be shifted by adding to S a local counterterm $\int d^2x\, f(\lambda(x))R_2(x)$ dependent on an arbitrary function $f(\lambda)$. In the context of nonlinear sigma models such term can be fixed by target space diffeomorphism invariance. For the model at hand this gives $w_\lambda = (8\pi\lambda)^{-1}$.

We see from (9.10) that while the theory has a vanishing beta function, its c-function: $c = 12\pi C(\lambda)$ has a non-trivial derivative with respect to the modulus λ. We can further observe that it is the broken global conformal symmetry that is responsible for the breakdown of gradient property. The stress–energy tensor on a flat surface is

$$T_{\mu\nu} = \frac{\lambda}{4\pi}\left(:\partial_\mu X\partial_\nu X: - \frac{\delta_{\mu\nu}}{2} :\partial_\gamma X\partial^\gamma X: \right) + \frac{Q}{4\pi}(\delta_{\mu\nu}\partial_\lambda\partial^\lambda - \partial_\mu\partial_\nu)X. \tag{9.13}$$

It has exactly the same form as the background charge model [26] with imaginary background charge. Note that in our theory there is no background charge. Moreover, since our theory is defined on a topological plane, the field X can be taken to be compact with an arbitrary radius. The correlation function

$$\langle T(z) J_{\lambda,z}(0)\rangle = -\frac{Q^2}{4\pi\lambda^2}\frac{1}{z^3} \tag{9.14}$$

means that special conformal transformations are broken by the boundary condition at infinity[10].

[10] The charge $\oint \underline{d}z\, z^2 T(z)$ does not vanish at infinity.

Another way to see the necessity to have theory defined on a sphere of large radius is in the context of nonlinear sigma model. There it is essential for the gradient formula to hold (at least in the leading order in the α' expansion) that the zero mode measure includes the dilaton contribution corresponding to spherical topology [19].

9.3. Bare gradient formula

Here we will show how the Wess–Zumino consistency condition for the local renormalization operation can be used to derive a different gradient formula. The main quantities in the new gradient formula are constructed using the anomalous contact terms present in $\mathcal{D}$ rather than correlation functions at finite separation. For this reason we call it a bare gradient formula. As a consequence of that the terms in that formula are defined modulo contact term ambiguities discussed in section 9.1. The new formula also suffers from potential infrared divergences in the metric. In this section however for the sake of brevity we will not introduce an explicit infrared cutoff and our manipulations with integrals will be formal. It is straightforward however to introduce such a cutoff with the main result correct up to some error terms vanishing when the cutoff is removed.

Using (7.49) the Wess–Zumino consistency condition

$$[(D(x_2) - \mathcal{D}(x_2)), (D(x_1) - \mathcal{D}(x_1))]\,\rangle\rangle = 0 \tag{9.15}$$

can be rewritten as[11]

$$\begin{aligned}&[[\Theta(x_2), \mathcal{D}(x_1)] - [\Theta(x_1), \mathcal{D}(x_2)] - [\beta(x_2), \mathcal{D}(x_1)]\\ &\quad+\beta(x_1)D(x_2) - \mathcal{D}(x_2)\Theta(x_1) + \mathcal{D}(x_1)D(x_2)]\,\rangle\rangle = 0.\end{aligned} \tag{9.16}$$

Applying to the above equation $\langle\langle\phi_i(y)$ on the left-hand side and integrating over x_1 we obtain

$$\begin{aligned}&\langle\langle\phi_i(y)[\mathcal{D}\Theta(x_2) - \beta^j\mathcal{D}\phi_j]\rangle\rangle + \langle\langle\phi_i(y)\int \mathrm{d}^2x_1\, \beta(x_1)D(x_2)\rangle\rangle + \langle\langle\mathcal{D}\phi_i(y)D(x_2)\rangle\rangle\\ &\quad-\mu\frac{\partial}{\partial\mu}(\langle D(x_2)\phi_i(y)\rangle_c - \delta^2(y-x_2)\partial_i\beta^j\langle\phi_j\rangle) = 0\end{aligned} \tag{9.17}$$

where we used the identities

$$\int dx_1[\Theta(x_1), \mathcal{D}(x_2)] = 0, \qquad \langle\langle\phi_i(y)\int dx_1[\beta^j(x_2), \mathcal{D}(x_1)]\phi_j(x_2)\rangle\rangle = 0. \tag{9.18}$$

As we know from section 7.3 $\mathcal{D}\Theta - \beta^j\mathcal{D}\phi_j$ is a pure-contact operation. Its field part $\beta^j\partial_\mu J_i^\mu - \partial_\mu J^\mu$ vanishes (is pure contact). Equation (9.17) expresses the contact terms with $\phi_i(y)$ via the operation $\mathcal{D}$. Integrating the above formula over x_2 with the weight $(x_2-y)^2$ and using

$$\mu\frac{\partial}{\partial\mu}\int \mathrm{d}^2x_2\,\langle D(x_2)\phi_i(y)\rangle_c(x_2-y)^2 = 0 \tag{9.19}$$

we obtain[12]

$$\partial_i\langle C_2\rangle = -H_{ij}\beta^j + \mathcal{L}_\beta W_i + Q_i \tag{9.20}$$

[11] Note that this form of the Wess–Zumino condition is linear in $\mathcal{D}$. This leads to essential simplifications in computations and also ensures that terms with tensorial sources in $\mathcal{D}$ do not contribute to the final gradient formula.
[12] Recall that the currents J_i^μ and the metric G_{ij} in (5.15) are ordinary operations so that $\langle\partial_i J_j^\mu\rangle = 0$.

where

$$H_{ij} = -G_{ij} - \frac{1}{4}\int d^2y\, y^2\big[\langle\partial_\mu J^\mu_j(0)\phi_i(y)\rangle + \langle\partial_\mu J^\mu_i(0)\phi_j(y)\rangle\big],$$

$$G_{ij} = -\frac{1}{4}\int d^2y\, y^2\langle\langle[\phi_i(0), \mathcal{D}\phi_j(y)]\rangle\rangle,$$

$$W_i = \frac{1}{4}\int d^2y\, y^2\langle D(y)\phi_i(0)\rangle_c,$$

$$Q_i = \frac{1}{4}\int d^2y\, y^2\langle\partial_\mu J^\mu_i(y)\Theta(0)\rangle_c. \tag{9.21}$$

Note that the tensor G_{ij} is symmetric. This follows from the fact that operations $\phi_i(y)$, $\phi_j(x)$ commute. The metric tensor H_{ij} can be also written in terms of integrated correlation functions:

$$H_{ij} = \frac{1}{4}\int d^2y y^2\left[\int d^2x\langle D(x)\phi_i(y)\phi_j(0)\rangle_c - \partial_i\beta^k\langle\phi_k(y)\phi_j(0)\rangle_c - \partial_j\beta^k\langle\phi_i(y)\phi_k(0)\rangle_c\right]. \tag{9.22}$$

According to our main infrared assumption (7.26) Q_i vanishes and we have a gradient formula

$$\partial_i c^{(0)} + g^{(0)}_{ij}\beta^j + b^{(0)}_{ij}\beta^j = 0 \tag{9.23}$$

where

$$c^{(0)} = \langle C_2\rangle - W_i\beta^i, \qquad g^{(0)}_{ij} = H_{ij}, \qquad b^{(0)}_{ij} = \partial_i W_j - \partial_j W_i. \tag{9.24}$$

The metric H_{ij} potentially suffers from the same infrared divergences as the correction to Zamolodchikov's metric defined in (8.15). We define the finite quantity entering (9.23) by subtracting these divergences.

When a loose power counting applies, the above quantities can be computed more explicitly using (5.15). In this case we have

$$H_{ij} = G_{ij} - \frac{1}{4}\int d^2y\, y^2\big[\langle\partial_\mu J^\mu_i(y)\phi_j(0)\rangle + \langle\partial_\mu J^\mu_j(y)\phi_i(0)\rangle\big], \tag{9.25}$$

$\langle C_2\rangle = \langle C\rangle$ and G_{ij}, W_i coincide with the respective quantities defined in formula (5.15). In the case when the currents J^μ_i are absent, formula (9.23) matches with the one obtained by Osborn [19].

9.4. Dressing transformations

For any gradient formula

$$\partial_i c + g_{ij}\beta^j + b_{ij}\beta^j = 0 \tag{9.26}$$

with a symmetric tensor g_{ij} and an antisymmetric tensor b_{ij} one can redefine c, b_{ij} and g_{ij} as

$$\tilde{c} = c + \beta^i c_{ij}\beta^j,$$

$$\tilde{g}_{ij} = g_{ij} - \mathcal{L}_\beta c_{ij},$$

$$\tilde{b}_{ij} = b_{ij} - (di_\beta c)_{ij}, \tag{9.27}$$

so that a gradient formula $\partial_i \tilde{c} = \tilde{g}_{ij}\beta^j + \tilde{b}_{ij}\beta^j$ holds. The tensor c_{ij} above is any tensor on the space of couplings that may depend on the couplings and the renormalization scale μ. We will refer to redefinitions (9.27) as dressing transformations. One can show that formula (8.14) is related to formula (9.23) by means of a dressing transformation specified by

$$c_{ij}^{\Lambda} = \int \mathrm{d}^2x\, G_{\Lambda}(x)\langle\phi_i(x)\phi_j(0)\rangle_c \tag{9.28}$$

so that

$$c = c^{(0)} - \beta^i c_{ij}^{\Lambda}\beta^j. \tag{9.29}$$

It is not hard to construct using dressing transformations a class of c-functions that monotonically decrease under the RG flow. Such functions c^f can be defined as

$$c^f = -3\pi \int \mathrm{d}^2x\, x^2 f(x^2)\langle\Theta(x)\Theta(0)\rangle_c, \tag{9.30}$$

where $f(x^2)$ is a function such that $f(0) = 1$, $f(x^2)$ decreases fast at infinity[13] and

$$x^{\mu}\partial_{\mu} f(x^2) < 0. \tag{9.31}$$

These potential functions satisfy a gradient formula

$$\partial_i c^f = -\big(g_{ij}^f + \Delta g_{ij}^f + b_{ij}^f\big)\beta^j, \tag{9.32}$$

where

$$g_{ij}^f = -3\pi \int \mathrm{d}^2x\, x^2[x^{\mu}\partial_{\mu} f(x^2)]\langle\phi_i(x)\phi_j(0)\rangle_c \tag{9.33}$$

$$\Delta g_{ij}^f = 3\pi \int \mathrm{d}^2x\, x^2[f(x^2) - 1]\big(\big\langle\partial_{\mu}J_i^{\mu}(x)\phi_j(0)\big\rangle + \big\langle\partial_{\mu}J_j^{\mu}(x)\phi_i(0)\big\rangle\big) \tag{9.34}$$

$$b_{ij}^f = \partial_i w_j^f - \partial_j w_i^f \tag{9.35}$$

$$w_i^f = 3\pi \int \mathrm{d}^2x\, x^2 f(x^2)\langle\phi_i(x)\Theta(0)\rangle_c. \tag{9.36}$$

Such smeared c-functions were first considered in [20].

9.5. Renormalization group transformation as a flow of couplings

As one can observe from the form of Callan–Symanzik equations (6.11), the scale transformation of correlation functions

$$\langle\phi_{i_1}(x_1)\phi_{i_2}(x_2)\ldots\Theta(y_1)\Theta(y_2)\ldots\rangle_c$$

even at finite separation is not fully compensated by the change of couplings λ^i. In addition to changing the couplings according to their beta functions and rotating the fields ϕ_i by the anomalous dimension matrices $\partial_i\beta^j$ the operators $\phi_i(x)$ and $\Theta(y)$ each shift by an additional total derivative: $\partial_{\mu}J_i^{\mu}(x)$ and $\partial_{\mu}J^{\mu}(y)$, respectively. If the currents J_i^{μ}, J^{μ} are not conserved, these shifts affect the scale transformation of the correlation functions taken at finite separation. This signals that more couplings need to be introduced to parameterize such additional terms in the Callan–Symanzik equations. Thus, to account for the current $J^{\mu}(y)$ it is customary to introduce dilaton couplings λ_D^i that couple to $\phi_i(x)\mu^2 R_2(x)$ terms in the Lagrangian[14]. The

[13] An exponential decrease would suffice for all purposes.
[14] A completeness of the set ϕ_i is assumed here as discussed in section 5.

generating functional Z depends on these couplings according to the functional differential equation

$$\frac{\delta \ln Z}{\delta \lambda_D^i(x)} = \frac{1}{2}\mu^2 R_2(x)\frac{\delta \ln Z}{\delta \lambda^i(x)}. \tag{9.37}$$

The introduction of this new set of couplings is natural if one bears in mind that coupling constant redefinitions are responsible for having different RG schemes. To renormalize a theory on a curved space one needs counterterms of the form $\phi_i(x)\mu^2 R_2(x)$. As usual such counterterms are defined up to arbitrary finite parts. Changing the dilaton couplings λ_D^i accounts for changing the finite parts in such counterterms. (Previously we assumed that such counterterms are fixed somehow which amounts to partially fixing the RG scheme. This resulted in the extra terms in the Callan–Symanzik equations.) Expanding the operator $C(x)$ in (5.15) as $C(x) = \beta_D^i \phi_i(x)$ we see that the coefficients β_D^i can now be naturally interpreted as the beta functions for the dilaton couplings. For the loose power counting case the Callan–Symanzik equation for correlators of stress–energy tensor takes the form

$$\mu\frac{\partial}{\partial\mu}\langle T_{\mu\nu}(y_1)T_{\alpha\beta}(y_2)\ldots\rangle_c = \beta^i\frac{\partial}{\partial\lambda^i}\langle T_{\mu\nu}(y_1)T_{\alpha\beta}(y_2)\ldots\rangle_c + \left\langle\Gamma^C_{\mu\nu}(y_1)T_{\alpha\beta}(y_2)\ldots\right\rangle_c$$
$$+\left\langle T_{\mu\nu}(y_1)\Gamma^C_{\alpha\beta}(y_2)\ldots\right\rangle_c + \ldots = \left(\beta^i\frac{\partial}{\partial\lambda^i} + \beta_D^i\frac{\partial}{\partial\lambda_D^i}\right)\langle T_{\mu\nu}(y_1)T_{\alpha\beta}(y_2)\ldots\rangle_c \tag{9.38}$$

where

$$\Gamma^C_{\mu\nu}(x) = (\partial_\mu\partial_\nu - g_{\mu\nu}\partial_\alpha\partial^\alpha)C(x). \tag{9.39}$$

We used (9.37) and (5.15) to obtain the last equality in (9.38). We see that the dilaton couplings account for mixings of the stress–energy tensor with trivially conserved currents $\Gamma^C_{\mu\nu}(x)$. With the enlarged set of couplings (λ, λ_D) the change in scale for correlators of stress–energy tensor components (at finite separation) is exactly compensated by the change in coupling constants. In particular for the c-function (2.10) we have

$$\mu\frac{\partial c}{\partial\mu} = \left(\beta^i\frac{\partial}{\partial\lambda^i} + \beta_D^i\frac{\partial}{\partial\lambda_D^i}\right)c. \tag{9.40}$$

We can also compute the derivatives of the c-function (2.10) with respect to the dilaton couplings. Using (3.3), (9.37) and the identity

$$\frac{\partial}{\partial\lambda_D^i} = \int \mathrm{d}^2x\,\frac{\delta}{\delta\lambda_D^i(x)} \tag{9.41}$$

we obtain

$$\frac{\partial c}{\partial\lambda_D^i} = -\frac{\partial}{\partial\lambda_D^i}\int \mathrm{d}^2x\, G_\Lambda(x)\langle\Theta(x)\Theta(0)\rangle_c = 2\int \mathrm{d}^2x\, G_\Lambda(x)\langle\Theta(x)\partial_\mu\partial^\mu\phi_i(0)\rangle_c. \tag{9.42}$$

Integrating by parts in (9.42), using

$$\mu\frac{\partial}{\partial\mu}\langle\phi_i\rangle = \beta^j\partial_j\langle\phi_i\rangle \tag{9.43}$$

and the assumption that

$$\int \mathrm{d}^2x\,\langle\phi_j(x)\phi_i(0)\rangle = \partial_j\langle\phi_i\rangle = \partial_i\langle\phi_j\rangle < \infty, \tag{9.44}$$

we obtain

$$\frac{\partial c}{\partial\lambda_D^i} = -g^D_{ij}\beta^j \tag{9.45}$$

where

$$g_{ij}^{D} = 2\int d^2x\,[G_0(x) - G_\Lambda(x)]\langle\phi_j(x)\partial_\mu\partial^\mu\phi_i(0)\rangle_c \tag{9.46}$$

is a symmetric tensor.

We can further show that the contraction of gradient formula (8.14) with the beta functions β^i gives Zamolodchikov's formula (3.1). This boils down to the identity

$$\beta_D^i\frac{\partial c}{\partial\lambda_D^i} = \beta^i\Delta g_{ij}\beta^j. \tag{9.47}$$

Using equations (9.45), (9.46) the left-hand side of equation (9.47) can be written as

$$\beta_D^i\frac{\partial c}{\partial\lambda_D^i} = 2\int d^2x\,[G_\Lambda(x) - G_0(x)]\langle\Theta(x)\partial_\mu\partial^\mu C(0)\rangle_c \tag{9.48}$$

while for the right-hand side we have

$$\beta^i\Delta g_{ij}\beta^j = 2\int d^2x\,[G_\Lambda(x) - G_0(x)]\big\langle\Theta(x)\beta^i\partial_\mu J_i^\mu(0)\big\rangle_c. \tag{9.49}$$

The last expression coincides with (9.48) by virtue of the identity $\partial_\mu\partial^\mu C(x) = \partial_\mu J_i^\mu(x)$ proven in section 7.3. This identity can be used because the two-point function in (9.49) is taken at finite separation. It is not hard to extend the proof of identity (9.47) to a more general case not assuming the loose power counting. Formula (9.47) shows in particular that the metric correction Δg_{ij} is necessary to account for the flow of dilaton coupling constants when the last ones are present.

The additional gradient formula (9.45) together with the main formula (8.14) implies that the c-function is stationary with respect to the couplings (λ, λ_D) at fixed points $\beta^i = 0$. The inverse follows from Zamolodchikov's formula (3.1) combined with formula (9.40). Thus, under our main set of assumptions and with a loose power counting the stationary points of the c-function are in a one-to-one correspondence with the fixed points.

10. Final comments

As we said in the introduction one of the motivations to obtain a general gradient formula came from string theory. In regard to potential applications of our result to the problem of constructing string effective actions it should be stressed that we worked throughout with normalized connected correlation functions while it is the unnormalized and disconnected ones which are relevant to string theory. This fact also explains why our results *seem* to be at odds with the conclusion of [15] that the Zamolodchikov c-function does not give a suitable string effective action. In the unnormalized correlators the dilaton zero mode ϕ_0 contributes an overall factor $e^{-2\phi_0}$ which results in having the same factor in c. Thus, stationarity of c with respect to ϕ_0 implies that c has to vanish at stationary points. This factor and the related problem disappear when one builds c out of normalized correlators as we do in this paper.

The aforementioned problem with c prompted various authors to switch to using what we call the bare gradient formula which was discussed in section 9.3. The negative side of this is that the metric that appears in that formula, being built from contact terms, does not have any positivity properties.

In the present paper we focused on a formal derivation of the new gradient formula and discussed its general properties. It would be instructive to illustrate how it works on concrete examples in conformal perturbation theory and nonlinear sigma models. We are planning to do this in a separate publication [27]. It is also interesting to understand better the implications of the new formula for string theory. We leave this question to future studies.

Acknowledgments

The work of DF was supported by the Rutgers New High Energy Theory Center. Both authors acknowledge the support of Edinburgh Mathematical Society.

References

[1] Schwinger J 1951 *Phys. Rev.* **82** 914
Schwinger J 1953 *Phys. Rev.* **91** 713
[2] Wallace D J and Zia R K P 1974 Gradient flows and the renormalization group *Phys. Lett.* A 325
[3] Wallace D J and Zia R K P 1975 Gradient properties of the renormalization group equations in multicomponent systems *Annu. Phys.* **92** 142
[4] Friedan D 1980 Nonlinear models in two epsilon dimensions *Phys. Rev. Lett.* **45** 1057
[5] Friedan D 1980 Nonlinear models in two + epsilon dimensions *PhD Thesis* University of California at Berkeley, published as Lawrence Berkeley Laboratory Publication LBL-11517
Friedan D 1985 *Annu. Phys., NY* **163** 318
[6] Fradkin E S and Tseytlin A A 1985 Effective field theory from quantized strings *Phys. Lett.* B **158** 316
[7] Fradkin E S and Tseytlin A A 1985 Quantum string theory effective action *Nucl. Phys.* B **261** 1
[8] Callan C G, Friedan D, Martinec E J and Perry M J 1985 Strings in background fields *Nucl. Phys.* B **262** 593
[9] Tseytlin A A 1986 Vector field effective action in the open superstring theory *Nucl. Phys.* B **276** 391
[10] Callan C G, Klebanov I R and Perry M J 1986 Quantum string theory effective action *Nucl. Phys.* B **278** 78
[11] Zamolodchikov A 1986 Irreversibility of the flux of the renormalization group in a 2-D field theory *JETP Lett.* **43** 730
[12] Tseytlin A A 1986 Conformal anomaly in a two-dimensional sigma model on a curved background and strings *Phys. Lett.* B **178** 34
[13] Zamolodchikov A 1987 Renormalization group and perturbation theory about fixed points in two-dimensional theory *Yad. Fiz.* **46** 1819
Zamolodchikov A 1987 *Sov. J. Nucl. Phys.* **46** 1090
[14] Tseytlin A A 1987 Sigma-model Weyl invariance conditions and string equations of motion *Nucl. Phys.* B **294** 383
[15] Tseytlin A A 1987 Conditions of Weyl invariance of the two-dimensional sigma model from equations of stationarity of the 'central charge' action *Phys. Lett.* B **194** 63
[16] Curci G and Paffuti G 1987 Consistency between the string background field equation of motion and the vanishing of the conformal anomaly *Nucl. Phys.* B **286** 399
[17] Osborn H 1988 String theory effective actions from bosonic sigma models *Nucl. Phys.* B **308** 629
[18] Osborn H 1990 General bosonic sigma models and string effective actions *Annu. Phys.* **200** 1
[19] Osborn H 1991 Weyl consistency conditions and a local renormalization group equation for general renormalisable field theories *Nucl. Phys.* B **363** 486
[20] Cappelli A, Friedan D and Latorre J I 1991 C theorem and spectral representation *Nucl. Phys.* B **352** 616
[21] Affleck I and Ludwig A 1991 Universal noninteger 'ground state degeneracy' in critical quantum systems *Phys. Rev. Lett.* **67** 161
[22] Friedan D and Konechny A 2004 On the boundary entropy of one-dimensional quantum systems at low temperature *Phys. Rev. Lett.* **93** 030402 (hep-th/0312197)
[23] Freedman D Z, Headrick M and Lawrence A 2006 On closed string tachyon dynamics *Phys. Rev.* D **73** 066015 (hep-th/0510126)
[24] Elitzur S 1983 The applicability of perturbation expansion to two-dimensional Goldstone systems *Nucl. Phys.* B **212** 501
[25] David F 1980 *Phys. Lett.* **96B** 371
David F 1981 *Commun. Math. Phys.* **81** 149
[26] Dotsenko Vl S and Fateev V A 1984 Conformal algebra and multipoint correlation functions in 2D statistical models *Nucl. Phys.* B **240** 312
[27] Friedan D and Konechny A 2010 (in preparation)