THE AREA LAW AS A CONSEQUENCE OF THE LOOP WAVE EQUATION

M. LÜSCHER
Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany

Received 11 December 1979

An argument is presented showing that the Wilson loop expectation value decays according to the area law, if it is assumed to satisfy the free loop wave equation.

1. Recently, Nambu [1] proposed that the Schwinger–Dyson equations satisfied by the (suitably renormalized) Wilson loop expectation in a pure non-abelian gauge theory

$$\psi(\mathcal{C}) = \left\langle \mathrm{Tr} \left\{ P\, exp \oint^{\mathcal{C}} \mathrm{d}x_\mu A_\mu \right\} \right\rangle , \tag{1}$$

P: path ordering, could perhaps be replaced, as a zeroth approximation, by the free loop wave equation

$$\{\delta^2/\delta x_\mu(\sigma)\,\delta x_\mu(\sigma) - M^4 x'(\sigma)^2\}\ \psi(\mathcal{C}) = 0\,,$$

$$x'_\mu = \mathrm{d}x_\mu/\mathrm{d}\sigma\,. \tag{2}$$

Here, $\mathcal{C}$ denotes a closed, infinitely differentiable curve in euclidean space–time $\mathbf{R}^d$, $x_\mu(\sigma)$ $(0 \leqslant \sigma \leqslant 2\pi)$ is a parametrization of $\mathcal{C}$ and M a constant mass parameter. Eq. (2) is expected to hold for double point free curves only. The loop wave equation has been studied in the context of the relativistic string [2], but the wave functions constructed there do not actually solve the Minkowski space version of eq. (2) and are therefore not immediately useful here. In this letter we show that eq. (2) together with some plausible "technical" assumptions imply that

$$\ln \psi(\lambda\, \mathcal{C}) \underset{\lambda\to\infty}{=} -\lambda^2 M^2 A(\mathcal{C}) + \text{lower orders}\,, \tag{3}$$

where $\lambda\, \mathcal{C}$ denotes the curve $\mathcal{C}$ scaled by a factor of λ and $A(\mathcal{C})$ is the area of a minimal surface bounded by $\mathcal{C}$ (for an introduction to the mathematics of minimal surfaces see ref. [3]). In particular it follows from this result that static quarks are confined by an asymptotically linearly rising potential.

2. One of the "technical" assumptions needed to prove eq. (3) is that $\psi(\mathcal{C})$ decays in a smooth fashion for large loops, i.e.

$$\ln \psi(\lambda \mathcal{C}) \underset{\lambda\to\infty}{=} -\nu(\lambda)\, W(\mathcal{C}) + \text{lower orders}\,, \tag{4}$$

where $\nu(\lambda)$ diverges for $\lambda \to \infty$ and $W > 0$. Inserting eq. (4) into eq. (2) and keeping the leading terms only yields

$$\nu(\lambda)^2 (\delta W/\delta x_\mu)^2 = (\lambda M)^4 x'^2\,. \tag{5}$$

It follows that $\nu(\lambda)$ is proportional to λ^2, i.e.

$$\nu(\lambda) = \lambda^2 M^2\,, \tag{6}$$

if W is suitably normalized. Thus. eq. (5) reduces to

$$(\delta W/\delta x_\mu)^2 = x'^2 \tag{7}$$

and since W is a reparametrization invariant functional of $x_\mu(\sigma)$, we also have

$$x'_\mu\, \delta W/\delta x_\mu = 0\,. \tag{8}$$

The differential equations (7) and (8) have many solutions so that we cannot immediately identify $W(\mathcal{C})$ with $A(\mathcal{C})$. The additional pieces of information used to establish the equality are positivity and scale covariance of W (a consequence of eqs. (4) and (6)):

$$W \geqslant 0\,, \quad W(\lambda \mathcal{C}) = \lambda^2 W(\mathcal{C})\,. \tag{9}$$

3. Before proceeding with the argument it is helpful to consider the analogous case of the Klein–Gordon equation

$$\{\partial^2/\partial x_\mu \partial x_\mu - M^2\}\ \psi(x) = 0, \quad (x \neq 0)\,. \tag{10}$$

The corresponding W-function is defined for $x \neq 0$ and satisfies

$$(\partial W/\partial x_\mu)^2 = 1, \quad W \geqslant 0, \quad W(\lambda x) = \lambda W(x)\,. \tag{11}$$

To solve the differential equation, we follow the method of characteristics (see e.g., ref. [4, pp. 90 ff.]). Thus, one first determines the characteristic curves in phase space as the solutions of

$$\frac{\mathrm{d}}{\mathrm{d}\tau} x_\mu = \frac{\partial H}{\partial p_\mu}\,, \quad \frac{\mathrm{d}}{\mathrm{d}\tau} p_\mu = -\frac{\partial H}{\partial x_\mu}\,, \quad p^2 = 1\,, \tag{12}$$

where $H \overset{\mathrm{def}}{=} \frac{1}{2}(p^2 - 1)$ is the Hamilton function associated with the differential equation (11). Explicitly,

$$x_\mu(\tau) = p_\mu \tau + a_\mu, \quad p_\mu, a_\mu = \text{constant}, \quad p^2 = 1. \tag{13}$$

The usefulness of characteristics stems from the fact that for any solution W of the differential equation (11) the equality

$$p_\mu = \partial W/\partial x_\mu\,, \tag{14}$$

holds along the whole characteristic (13), if it is true at one point (x, p). In particular, the change of W along these lines is given by

$$\mathrm{d}W/\mathrm{d}\tau = 1\,. \tag{15}$$

To have $W \geqslant 0$ it is therefore necessary that all points $(x_\mu, \partial W/\partial x_\mu)$ in phase space determine characteristics passing through the origin $x = 0$, where $\partial W/\partial x_\mu$ may be discontinuous and eq. (15) stops to hold. It follows that

$$\partial W/\partial x_\mu = x_\mu/|x|$$

and since $W \to 0$ as $x \to 0$ by scale covariance, we are left with

$$W = |x|\,. \tag{16}$$

4. In the strong case we have to solve many first order partial differential equations simultaneously. The associated Hamilton functions are

$$H = \int_0^{2\pi} \mathrm{d}\sigma\, \{\tfrac{1}{2} f(p^2 - x'^2) + g x' \cdot p\}\,, \tag{17}$$

where $f(\sigma)$ and $g(\sigma)$ are arbitrary smooth periodic functions. The equations for the characteristics thus read

$$\frac{\partial}{\partial\tau} x_\mu = f p_\mu + g x'_\mu\,, \quad \frac{\partial}{\partial\tau} p_\mu = [-f x'_\mu + g p_\mu]'\,,$$
$$p^2 = x'^2, \quad x' \cdot p = 0\,. \tag{18}$$

Of course, these are the familiar classical string equations [2], whose solutions $x_\mu(\tau, \sigma)$ describe minimal surfaces. If the functions f and g are varied, the solution of eq. (18) with prescribed initial data $x_\mu(0, \sigma)$, $p_\mu(0, \sigma)$ will change. The minimal surface Σ swept out, however, remains the same.

Suppose now that W solves eqs. (7) and (8) and has the additional properties (9). Each point $[x_\mu(\sigma), \delta W/\delta x_\mu(\sigma)]$ in phase space then determines a unique characteristic curve (minimal surface) Σ passing through it. W changes in a simple fashion along Σ: if a closed curve $\mathcal{C}$ on Σ is smoothly deformed into another curve $\mathcal{C}'$, we have $W(\mathcal{C}') - W(\mathcal{C})$ = (signed) area swept out during the deformation $\mathcal{C} \to \mathcal{C}'$.

In particular, deformations on one side of $\mathcal{C}$ will decrease W. Since $W \geqslant 0$, we conclude that $\mathcal{C}$ actually encloses a part of Σ. Furthermore, by scale covariance, $W(\mathcal{C}')$ approaches zero, when $\mathcal{C}'$ is contracted to a point. $W(\mathcal{C})$ is thus equal to the area of the piece of Σ, which is enclosed by $\mathcal{C}$. In other words $W(\mathcal{C}) = A(\mathcal{C})$, as asserted.

5. The proof of the area law given above is not mathematically rigorous. In particular, we implicitly assumed that W was continuously differentiable. There is little reason to do so in fact, even if the Wilson loop expectation $\psi(\mathcal{C})$ is differentiable. Suppose for example that

$$\psi(\mathcal{C}) = \sum_i \psi_i(\mathcal{C})\,,$$

$$\ln \psi_i(\lambda\mathcal{C}) \underset{\lambda\to\infty}{=} -\lambda^2 M^2 W_i(\mathcal{C}) + \text{lower orders}, \tag{19}$$

where the W_i's are analytic functionals. In this case,

$$W(\mathcal{C}) = \min\{W_i(\mathcal{C})\}\,, \tag{20}$$

will not in general be continuously differentiable. Also, it it not difficult to argue that $A(\mathcal{C})$ or its first derivative must be discontinuous at some curves $\mathcal{C}$, which bound several minimal surfaces.

As a working hypothesis, it is consistent to assume that all discontinuities of W arise from the presence of competing terms in the large loop expansion of ψ as explained above. Instead of continuous differentiability, one thus requires that

(i) W is piecewise (real) analytic.

(ii) (cf. eq. (20)) W can be analytically continued across the discontinuities. The analytic continuation $\widetilde{W}$ is bounded by W: $\widetilde{W}(\mathcal{C}) \geqslant W(\mathcal{C})$.

The argument of the preceding section can then be adjusted easily, so that one arrives at eq. (3) again.

6. It is to be expected that there are many solutions of the loop wave equation, which behave like $\psi(\mathcal{C})$ for uniformly large loops (eq. (3)). To completely fix the solution one would also have to specify the singularity structure at small and self-crossing loops. It seems reasonable (but somewhat arbitrary) to choose singularities appropriate for a string Green's function. In this case, $\psi(\mathcal{C})$ can be formally constructed from the loop heat kernel $K(z, x)$ [5]:

$$\psi(\mathcal{C}) \propto \int \mathcal{D}[z]\, K(z,x)\,, \tag{21a}$$

$$K(z,x) \propto \int\limits_{\phi_\mu(z(\sigma))=x_\mu(\sigma)} \mathcal{D}[\phi_\mu] \times \exp -\frac{M^2}{2}\int\limits_D d^2z\,(\partial_a\phi_\mu)^2\,. \tag{21b}$$

Here, $z_a(\sigma)$, $0 \leqslant \sigma \leqslant 2\pi$, $a = 1, 2$, denotes a closed curve in the plane $\mathbf{R}^2$ bounding a region D. The fields ϕ_μ integrated over in eq. (21b) are defined on D and have prescribed values on ∂D. Although the "proper coordinate" representation (21) is not very explicit, it is easy to check eq. (3). Noting

$$\psi(\lambda\mathcal{C}) \propto \int \mathcal{D}[z]\,.\int\limits_{\phi_\mu(z(\sigma))=x_\mu(\sigma)} \mathcal{D}[\phi_\mu] \times \exp -\lambda^2\frac{M^2}{2}\int\limits_D d^2z\,(\partial_a\phi_\mu)^2\,, \tag{22}$$

we see that an asymptotic expansion for $\lambda \to \infty$ can be obtained by the saddle point method. The absolute minimum of

$$\frac{1}{2}\int\limits_D d^2z\,(\partial_a\phi_\mu)^2, \quad \phi_\mu(z(\sigma)) = x_\mu(\sigma)\,,$$

when $\phi_\mu(z)$ and $z_a(\sigma)$ are allowed to vary, is known to be the least area enclosed by $\mathcal{C}$ [3], so that the leading term in the saddle point expansion matches with eq. (3).

I am indebted to P. Di Vecchia, P. Weisz and K. Symanzik for helpful discussions.

References

[1] Y. Nambu, Phys. Lett. 80B (1979) 372.

[2] C. Rebbi, Phys. Rep. 12C (1974) 1;
J. Scherk, Rev. Mod. Phys. 47 (1975) 123.

[3] R. Courant, Dirichlet's principle, conformal mapping, and minimal surfaces (Springer, Berlin, (1977), reprint;
F.J. Almgren Jr., Plateau's Problem (Benjamin, New York, 1966).

[4] S. Flügge, Handbuch der Physik, Vol. I (Springer, Berlin, 1956).

[5] A.M. Polyakov, Gauge fields as rings of glue, preprint (1979).