Vortex Duality in Higher Dimensions

Beekman, A.J.

Chapter 6

Emergent gauge symmetry and duality

Breaking symmetry is easy but making symmetry is hard: this wisdom applies to global symmetry but not to local symmetry. The study of systems controlled by emergent gauge symmetry has become a mainstream in modern condensed matter physics. Although one discerns as a fundamental gauge symmetry only electromagnetism in the ultraviolet of condensed matter physics, it is now very well understood that in a variety of circumstances gauge symmetries that do not exist on the microscopic scale control the highly collective physics on the macroscopic scale. An intriguing but unresolved issue is whether the gauge structures involved in the Standard Model of high energy physics and perhaps even general relativity could be of such an emergent kind.

Up to now we have not focussed much on the emergence of gauge symmetries—rather we have taken them for granted as a either an unrelated coincidence or as a logical but still auxiliary tool in the vortex duality. This chapter discusses some of the deeper, underlying gauge principles that not only facilitate understanding the nature of the disordering transition, but even provide a new viewpoint to gauge symmetry in general, possibly adding to our comprehension of its importance.

A gauge symmetry is said to be emergent when it is not present in the microscopic model of the constituent particles or fields, but arises in the effective theory as a collective degree of freedom. We have encountered two examples in this work:

1. the "stay-at-home" gauge invariance associated with (doped) Mott insulators, expressing a local conservation law (see §2.3.4);

2. the global-to-local symmetry correspondence in the strong/weak (i.e. Kramers–Wannier, S-) dualities, or the expression of the Goldstone mode as a dual gauge field (see §§2.4.2, 3.1).

In the common perception these appear as quite different. Here we clarify that at least in the context of bosonic physics they are actually closely related. In fact, these highlight complementary aspects of the vacuum structure, and it just depends on whether one views the vacuum either using the canonical/Hamiltonian language (stay-at-home) or field-theoretical/Lagrangian (local–global duality) language.

In this chapter we shall first go through the Bose-Hubbard model/vortex–boson duality again in §6.1, emphasizing the dual aspects of the emerging gauge symmetries. The ability to switch back and forth between Hamiltonian and Lagrangian viewpoints yields some entertaining vistas on this well-understood theory. In particular, the condensate of vortices is to be understood as a coherent superposition of all possible vortex configurations, and we will show that this is completely equivalent as adding gauge symmetry to phase correlations.

To make the case that it can yield new insight, we apply it in §6.2 to the less familiar context of dualities in quantum elasticity. This deals with the description of quantum liquid crystals in terms of dual condensates formed from the translational topological defects (dislocations) associated with the fully ordered crystal. Using the Lagrangian language it was argued that such quantum nematics are equivalent to (linearized) Einstein gravity [43]. Here we will demonstrate that this is indeed controlled by the local symmetry associated with linearized gravity: translations are gauged, turning into infinitesimal Einstein transformations.

6.1 Vortex duality versus Bose-Mott insulators

A mainstream of the gauge theories in condensed matter physics dates back to the late 1980s when the community was struggling with the fundamentals of the problem of high-T_c superconductivity. It was recognized early on that this has to do with doping the parent Mott insulators and this revived the interest in the physics of the Mott insulating state itself [58, 92–94]. The point of departure is the Hubbard model for electrons,

$$H_{\mathrm{FH}} = -t \sum_{<ij>\sigma} (\hat{c}^{\dagger}_{i\sigma}\hat{c}_{j\sigma} + \hat{c}^{\dagger}_{j\sigma}\hat{c}_{i\sigma}) + U\sum_{i} \hat{n}_{i\uparrow}\hat{n}_{i\downarrow}\,, \tag{6.1}$$

describing fermions $\hat{c}^{\dagger}_{i\sigma}$ on site i with spin σ, hopping on a lattice with rate t, subjected to a strong local Coulomb interaction U. Here $\hat{n}_{i\sigma} = \hat{c}^{\dagger}_{i\sigma}\hat{c}_{i\sigma}$ is the fermion number operator. A much simpler problem is the Bose-Hubbard model of §2.3. It describes spinless bosons created by $\hat{b}^{\dagger}_{i}$ hopping on a lattice with a rate t subjected to an on-site repulsion U,

$$H_{\mathrm{BH}} = -t \sum_{<ij>} \hat{b}^{\dagger}_{i}\hat{b}_{j} + U\sum_{i} \hat{n}^{2}_{i}\,. \tag{6.2}$$

Again $\hat{n}_i = \hat{b}^{\dagger}_{i}\hat{b}_i$ is the boson number operator. We assume in the remainder that the system is at "zero chemical potential", meaning that on average there is an integer number of fermions or bosons n_0 per site.

6.1.1 Stay-at-home gauge symmetry

Now these models are invariant under a *global* symmetry,

$$\hat{c}^{\dagger}_{i\sigma} \to \hat{c}^{\dagger}_{i\sigma}\mathrm{e}^{\mathrm{i}\alpha_{\sigma}} \qquad \text{or} \qquad \hat{b}^{\dagger}_{i} \to \hat{b}^{\dagger}_{i}\mathrm{e}^{\mathrm{i}\alpha}, \tag{6.3}$$

where the symmetry transformation is a scalar variable α that is constant for all lattice sites. But in the limit $U/t \to \infty$ the hopping term vanishes, and this symmetry is promoted to a local symmetry,

$$\begin{aligned} \hat{c}^{\dagger}_{i\sigma} &\to \hat{c}^{\dagger}_{i\sigma}\mathrm{e}^{\mathrm{i}\alpha_{i\sigma}}\,, & & & \hat{b}^{\dagger}_{i} &\to \hat{b}^{\dagger}_{i}\mathrm{e}^{\mathrm{i}\alpha_{i}} \\ \hat{c}_{i\sigma} &\to \mathrm{e}^{-\mathrm{i}\alpha_{i\sigma}}\hat{c}_{i\sigma}\,, & &\text{or} & \hat{b}_{i} &\to \mathrm{e}^{-\mathrm{i}\alpha_{i\sigma}}\hat{b}_{i}\,, \\ \hat{n}_i &= \sum_{\sigma}\hat{c}^{\dagger}_{i\sigma}\hat{c}_{i\sigma} \to \hat{n}_i & & & \hat{n}_i &= \hat{b}^{\dagger}_{i}\hat{b}_i \to \hat{n}_i. \end{aligned} \tag{6.4}$$

One discovers that a gauge symmetry emerges which controls the physics at long distances, while it is non-existent at the microscopic scale. This is the point of departure of a mainstream school of thought in condensed matter physics. In the fermionic model, there is still a dynamical spin system at work at low energies. Using various "slave-constructions" it was subsequently argued that quantum spin liquids characterized by fractionalized excitations can be realized when the resulting compact $U(1)$ gauge theory would end up in a deconfining regime. Conversely, the spinless Bose variety is completely featureless since it does not seem to break a manifest symmetry while low energy degrees of freedom are absent.

Note that at any finite U/t this gauge symmetry would be strictly broken; still at large values of this parameter it is a good idea to start from the gauge-invariant ground state, with deviation from this state entering as excitations.

The complete Hubbard models are defined in term of particle creation and annihilation operators, but in the Mott insulating state, the number of particles is *locally* conserved, i.e. conserved at each site separately, and only the number operator is present in the resulting Hamiltonian. The emergence of the gauge symmetry is caused by this local number conservation. One could picture that the annihilation–creation combination $c_i^\dagger c_i$ is now "tied" by emergent gauge bosons as force carriers: the particles are told to "stay at home". Indeed, the doped Mott insulator is described in such terms, leading to spin-charge separation and so forth.

This emergent gauge symmetry is not restricted to lattice models. Take for instance the effective Landau model describing the superfluid Eq. (2.1),

$$H = \int \mathrm{d}^3x \; \frac{1}{2}\tau|\nabla\Psi|^2 + \frac{1}{2}\alpha|\Psi|^2 + \frac{1}{4}\beta|\Psi|^4. \tag{6.5}$$

We have inserted a parameter τ for convenience. This model is invariant under global $U(1)$ symmetry $\Psi(x) \to \mathrm{e}^{\mathrm{i}\alpha}\Psi(x)$, where α is constant in space. But if we were to suppress the fluctuation term $\tau \to 0$, then this would be promoted to a local symmetry $\alpha \to \alpha(x)$. In other words, in the absence of fluctuations of the order parameter, the superfluid is indistinguishable from a superconductor. Furthermore the rigidity of the order parameter is now no longer enforced by a Goldstone mode, but by a local conservation law.

6.1.2 Vortex–boson duality

As detailed in §2.3, the Bose-Hubbard model at zero chemical potential can be mapped onto the XY- or phase-only model, which in turn maps onto the superfluid in the weak-coupling and continuum limit. We saw in §2.4 and chapter 3 that the phase transition to the Mott insulator is then formulated by the proliferation of topological defects, in this case vortices.

We needed to pass from the Goldstone field φ to its canonical conjugate, the supercurrent w_μ. This is the Noether current of the global symmetry $\varphi \to$

$\varphi + \varepsilon$. The smoothness of the Goldstone field ensures that the supercurrent is conserved $\partial_\mu w_\mu = 0$, which can be enforced by expressing it in terms of a dual gauge field,

$$w_\mu = \epsilon_{\mu\nu\lambda_1\cdots\lambda_{d-2}} \partial_\nu b_{\lambda_1\cdots\lambda_{d-2}}, \tag{6.6}$$

which is invariant under non-compact $U(1)$ gauge transformations,

$$b_{\lambda_1\cdots\lambda_{d-2}} \to b_{\lambda_1\cdots\lambda_{d-2}} + \partial_{[\lambda_1} \varepsilon_{\lambda_2\cdots\lambda_{d-2}]}, \tag{6.7}$$

where $\varepsilon_{\lambda_2\cdots\lambda_{d-2}}$ is any smooth $d-3$-form field. These gauge fields have the natural interpretation as the force carriers of the interactions between vortex excitations.

Once again, the global symmetry of the original model seems to be promoted to a local symmetry, but surely this non-compact symmetry of the Coulomb or superfluid phase is completely different from the compact $U(1)$ of the stay-at-home gauge invariance of the Higgs or Mott insulating phase.

The next step is to consider what happens across the phase transition. The vortices proliferate into a 'tangle of vortex world lines' or 'string foam', which is as a fluid medium minimally coupled to the dual gauge fields, which therefore undergo an Anderson–Higgs mechanism. The long-range correlations mediated by massless gauge fields now turn short-range.

6.1.3 The vortex condensate generates stay-at-home gauge

Up to this point we have just collected and reviewed some well-known results on phase dynamics. However, at first sight it might appear as if the matters discussed in the two previous subsections are completely unrelated. Departing from the Bose-Hubbard model the considerations of the previous subsection leave no doubt that in one or the other way the dual vortex '$d-2$-form superconductor' can be adiabatically continued all the way to the strongly coupled Bose-Mott insulator of the first subsection. The standard way to argue this is by referral to the excitation spectrum. The Bose-Mott insulator is characterized by a mass gap $\sim U$ (at strong coupling), and a doublet of "holon" (vacancy) and "doublon" (doubly-occupied site) propagating excitations being degenerate at zero chemical potential (see §§2.3,2.4.5 and 3.2). The vortex superconductor is a relativistic $U(1)/U(1)$ Higgs condensate characterized by a Higgs mass (a gap) above which one finds a doublet massive gauge bosons. In this regard there is a precise match. However, in the canonical formalism one also discovers the emergent $U(1)$ invariance associated with the sharp quantization of local number density in the Mott insulator. What has happened to this important symmetry principle in the vortex superconductor?

The answer is: the emergent compact $U(1)$ gauge symmetry of the Mott insulator is actually a generic part of the physics of the relativistic superconductor.

The argument is amazingly simple. The stay-at-home gauge does not show up explicitly in the Higgsed action describing the dual vortex condensate, for the elementary reason that all the quantities in this action are associated with the vortices which are in turn in a perfect non-local relation with the original phase variables. However, we know precisely what this dual superconductor is in terms of those phase variables. We can resort to a first quantized, world line description of the vortex superconductor, putting back "by hand" the phase variables. This constitutes a tangle of world lines of vortices, warping the original phases, and eventually we can even map that back to a first quantized wave function written as a coherent superposition of configurations of the phase field. To accomplish this in full one needs big computers [32, 33], but for the purposes of scale and symmetry analysis the outcomes are obvious.

The penetration depth λ_V of the dual vortex superconductor just coincides with the typical distance between vortices. At distances much shorter than λ_V the vortices do not scramble the relations between the phases at spatially separated points and at these scales the system behaves as the ordered superfluid,

$$\langle b^\dagger(r) b(0) \rangle \to \text{constant}, \quad r \ll \lambda_V, \tag{6.8}$$

However, at distances of order λ_V and larger, the vacuum turns into a coherent quantum superposition of "Schrödinger cat states" where there is either none, or one, or whatever number of vortices in between the two points 0 and r whose correlation of the phases of bosons we wish to know, see Fig. 6.1. We have arrived at exposing the simple principle which is the central result of this chapter: *since the vortex configurations are in coherent superposition, the phases acquire a full compact $U(1)$ gauge invariance.* Here is how to understand the physical concept: focus on the direction of the phase at the origin and look at the phase arrow at some distance point r. Consider a particular configuration of the vortices, and in this realization the distant phase

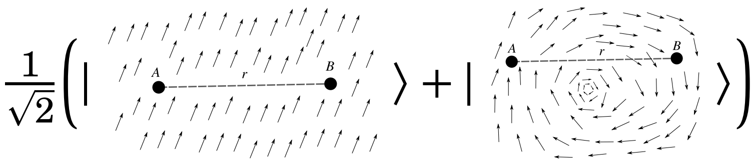


Figure 6.1: In the vortex condensate the correlation of the phase between a point A and another point B a distance r apart is in a superposition of having zero, one or any number of vortices in between. As such the phase at B with respect to that at A is completely undefined: it has acquired a full gauge invariance in the sense that any addition to the phase is an equally valid answer

will point in some definite direction which will be different from the phase at the origin as determined by the particular vortex configuration. However, since all different vortex configurations are in coherent superposition and therefore "equally true at the same time", all orientations of the phase at point r are also "equally true at the same time" and this is just the precise way to formulate that a compact $U(1)$ gauge symmetry associated with ϕ has emerged at distances λ_{V}.

The implication is that via Eq. (6.4) the emerging stay-at-home gauge invariance implies that in the Higgs condensate the number density associated with the bosons condensing in the dual superfluid becomes locally conserved on the scale λ_V. The Mottness therefore sets in only at scales larger than this λ_V. Notice that this mechanism does in fact not need a lattice: it is just a generic property of the field theory itself, which is independent of regularization. In fact, the seemingly all important role of the lattice in the standard reasoning in condensed matter when dealing with these issues is a bit of tunnel vision. It focusses on the strong-coupling limit where for large U, $\lambda_{\mathrm{V}} \to a$, the lattice constant. However, upon decreasing the coupling strength, the stay-at-home gauge emerges at an increasingly longer length scale λ_{V}, to eventually diverge at the quantum phase transition. Close to the quantum critical point the theory has essentially forgotten about the presence of the lattice, just remembering that it wants to conserve number locally which is the general criterion to call something an insulator. In fact, Mottness can exist without a lattice altogether. A relativistic superconductor living in a perfect 2+1d continuum is physically reasonable. Since duality works in both directions, this can be in turn viewed as a quantum disordered superfluid, where the number density associated with the bosons comprising the superfluid becomes locally conserved.

By inspecting closely this simple vortex duality we have discovered a principle which might be formulated in full generality as: *the coherent superposition of the disorder operators associated with the condensation of the disorder fields has the automatic consequence that the order fields acquire a gauge invariance associated with the local quantization of the operators conjugate to the operators condensing in the order field theory.* We suspect that this principle might be of use also in the context of dualities involving more complex field theories.

6.2 Quantum nematic crystals and emergent linearized gravity

To substantiate this claim, let us now inspect a more involved duality which is encountered in quantum elasticity, where the principle reveals the precise reasons for why quantum liquid crystals have dealings with general relativity. Einstein himself already suggested the metaphor that the spacetime of general relativity is like an elastic medium. Is there a more literal truth behind it? In recent years Hagen Kleinert has been forwarding the view that quite deep analogies exist between plastic media (solids with topological defects) and Einsteinian spacetime [41, 42]. There appears room for the possibility that at the Planck scale an exotic "solid" (the "world crystal") is present, turning after coarse graining into the spacetime that we experience.

It turns out that this subject matter has some bearing on a much more practical question: what is the general nature of the quantum hydrodynamics and rigidity of quantum liquid crystals? Quantum liquid crystals [82] are just the zero temperature versions of the classical liquid crystals found in computer displays. These are substances characterized by a partial breaking of spatial symmetries, while the zero temperature versions are at the same time quantum liquids. Very recently indications have been found for variety of such quantum liquid crystals in experiment [95–100]. In the present context we are especially interested in the "quantum smectics" and "quantum nematics" found in high-T_c cuprates [84, 96–98] which appear to be also superconductors at zero temperature. Such matter should be, at least in the long-wavelength limit, governed by a bosonic field theory, and this "theory of quantum elasticity" [40, 45, 83, 101] is characterized by dualities that are richer, but eventually closely related to the duality discussed in the previous section.

Departing from the quantum crystal, the topological agents which are responsible for the restoration of symmetry are the dislocations and disclinations. The disclinations restore the rotational symmetry and the topological criterion for liquid crystalline order is that these continue to be massive excitations. The dislocations restore translational symmetry, and these are in crucial regards similar to the vortices of the previous section. In direct analogy with the Mott insulator being a vortex superconductor, the superconducting smectics and nematics can be universally viewed as dual “stress superconductors” associated with Bose condensates of quantum dislocations.

Using the geometrical correspondences of Kleinert [41, 42], arguments were put forward suggesting that the Lorentz-invariant version of the superfluid nematic in 2+1d is characterized by a low energy dynamics that is the same as at least linearized gravity [43]. Very recently it was pointed out that this appears also to be the case in the 3+1d case [102, 103]. A caveat is that Lorentz invariance is badly broken in the liquid crystals as realized in condensed matter physics. This changes the rules drastically and although the consequences are well understood in 2+1d [40, 45, 83] it remains to be clarified what this means for the 3+1d condensed matter quantum liquid crystals. The currently unresolved issue is how the gravitons of the 3+1d relativistic case imprint on the collective modes of the non-relativistic systems.

Here we want to focus on perhaps the most fundamental question one can ask in this context: although general relativity is not a Yang–Mills theory, it is uniquely associated with the gauge symmetry of general covariance or diffeomorphisms. Quite generally, attempts to identify “analogue” or “emergent” gravity in condensed matter systems have been haunted by the problem that general covariance is quite unnatural in this context. The gravity analogues currently contemplated in condensed-matterlike systems usually get as far as to identify a non-trivial geometrical parallel transport of the matter, that occurs in a “fixed frame” or “preferred metric” [104–111]; in other cases this issue of the mechanism of emerging general covariance is simply not addressed [112–114]. As we shall discuss, crystals are manifestly non-diffeomorphic. However, the relativistic quantum nematics appear to be dynamically similar to Einsteinian spacetime. For this to be true, in one way or another general covariance has to emerge in such systems. How does this work?

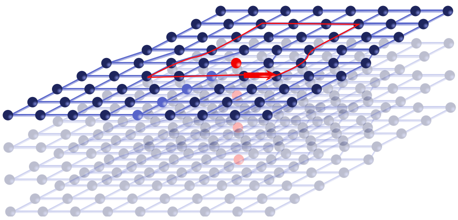
In close parallel with the vortex duality "toy model" of the previous section, we will explicitly demonstrate in this section that indeed general covariance is dynamically generated as an emergent IR symmetry. However, there is a glass ceiling: the geometry is only partially gauged. Only the infinitesimal "Einstein" translations fall prey to an emergent gauge invariance while the Lorentz transformations (rotations) remain in a fixed frame. This prohibits the inclusion of black holes and so forth, but this symmetry structure turns out to be coincident with the 'gauge fix' that is underlying *linearized* gravity. The conclusion is that relativistic quantum nematics constitute a medium that supports gravitons, but nothing else than gravitons.

For this demonstration we have to rely on the detour for the identification of the local symmetry generation as introduced in the previous section. Different from the Bose-Mott insulator, there is no formulation available for the quantum nematic in terms of a simple Hamiltonian where one can directly read off the equivalent of the stay-at-home gauge symmetry. We have therefore to find the origin of the gauging of the Einstein translations in the physics of the dislocation Bose condensate, but this will turn out to be a remarkably simple and elegant affair.

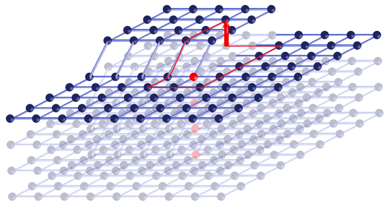
The remainder of this section is organized as the previous one. In section 6.2.1 we will first collect the various bits and pieces: a sketch of the way that "dislocation duality" associates the relativistic quantum nematic state with a crystal that is destroyed by a Bose condensate of dislocations. In section 6.2.2 we will subsequently review Kleinert's "dictionary" relating quantum elasticity and Einsteinian geometry, while at the end of this subsection we present the mechanism of gauging Einstein translations by the dislocation condensate. For simplicity we will focus on the 2+1d case; the generalities we address here apply equally well to the richer 3+1d case.

6.2.1 The quantum nematic as a dislocation condensate

Let us first introduce the field-theoretical side [40, 45, 83, 101]. The theory of quantum elasticity is just the 19th century theory of elasticity but now embedded in the Euclidean spacetime of thermal quantum field theory. To keep matters as simple as possible we limit ourselves to the Lorentz-invariant "world crystal", just amounting to the statement that we are dealing with a



(a) Edge dislocation



(b) Screw dislocation

Figure 6.2: Dislocation lines (red spheres) in the relativistic 3D "world crystal" (two space and one time direction), formed by insertion of a half-plane of particles. Shown in red is the contour that measures the mismatch quantized in the Burgers vector (red arrow). If the Burgers vector is orthogonal to the dislocation line it is an edge location; if the Burgers vector is parallel it is a screw dislocation. In non-relativistic 2+1d there are only edge dislocations, since the Burgers vector is always purely spatial.

2+1d elastic medium being isotropic, both in space and time directions,

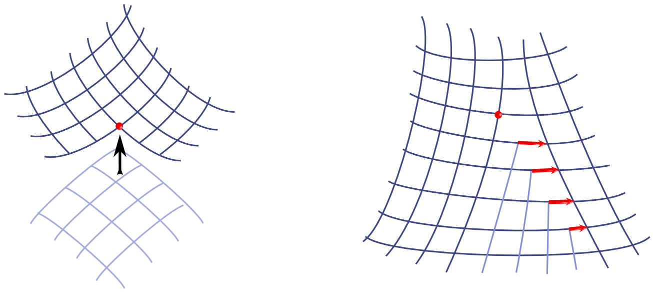
$$Z = \int Dw \; \mathrm{e}^{-S_{\mathrm{el}}} ,$$
$$S_{\mathrm{el}} = \int \mathrm{d}\tau \mathrm{d}x^2 \; \left[\mu w_{\mu\nu} w^{\mu\nu} + \frac{\lambda}{2} w^2_{\mu\mu} \right] , \tag{6.9}$$

where,

$$w_{\mu\nu} = \frac{1}{2} \left(\partial_\mu u_\nu + \partial_\nu u_\mu \right) , \tag{6.10}$$

are the strain fields associated with the displacements u_ν of the "world crystal atoms" relative to their equilibrium positions. Here μ and λ are the shear modulus and the Lamé constant of the world crystal, respectively. At first view this looks like a straightforward tensorial generalization of the scalar field theory of the previous section. For the construction of the nematics one can indeed think about the displacements as "scalar fields with flavours" since this only involves the "Abelian sector" of the theory associated with translations. One should keep in mind however that one is breaking Euclidean space down to a lattice subgroup and this is associated with non-Abelian, infinite and semi-direct symmetry structure: the full theory beyond the dislocation duality is a much more complicated affair.

These issues become manifest when considering the topological defects: the dislocations and disclinations. The dislocation is the topological defect



(a) Disclination (b) Stack of dislocations

Figure 6.3: (a) 90° disclination in a square lattice. A wedge is inserted into a cut in the lattice. There is now one lattice point with five instead of four neighbouring sites (red); going along a contour around this point will result in an additional 90° rotation. The associated topological charge is the Frank vector, orthogonal to the plane and of size 90°. As the dislocation, in 2+1d spacetime the disclination point will trace out a world line. (b) Disclination as a stack of dislocations. Hence a disclination corresponds to a uniform polarization of Burgers vectors. As long as disclinations are massive, e.g. in the quantum nematic, dislocations appear only with balanced opposite Burgers vectors.

associated with the restoration of the translations. The dislocation can be viewed as the insertion of a half-plane of extra atoms terminating at the dislocation core. One immediately infers that it carries a vectorial topological charge: the Burgers vector indexed according to the Miller indices of the crystal. In 2+1 dimensions the dislocation is a particle (like the vortex) and as an extra complication the Burgers vector can either lie perpendicular ["edge dislocation", Fig. 6.2(a)] or parallel ["screw dislocation", Fig. 6.2(b)] to the propagation direction of its world line. The disclination is on the other hand associated with the restoration of the rotational symmetry. This can be obtained by the Volterra construction: cut the solid, insert a wedge and glue together the sides [see Fig. 6.3(a)]. This carries a vectorial charge (the Frank vector) as well. Finally dislocations and disclinations are not independent. On the one hand, the disclination can be viewed as a stack of dislocations with parallel Burgers vectors [Fig. 6.3(b)], while the dislocation can be viewed as a disclination–antidisclination pair displaced by a lattice constant.

Dislocations and disclinations do however have a distinguishable identity and this enables a tight, topological definition of quantum smectic and nematic order. A state where dislocations have spontaneously proliferated and condensed, while the disclinations are still massive, is a quantum liquid crystal. Since a disclination is coincident with a "uniform magnetization" of Burgers vectors, one cannot have a net density of parallel Burgers vectors as long as disclinations are suppressed [see Fig. 6.3(b)]. The Burgers vectors of the dislocations in the condensate have to be anti-parallel and therefore the dislocation breaks orientations rather than rotations, with the ramification that the order parameter is a director instead of a vector.

Finally, when all orientations of the Burgers vectors are populated equally in the condensate, one deals with a nematic breaking only space rotations. When only a particular Burgers vector orientation is populated one is dealing with a smectic because the translations are only restored in the direction of the Burgers vector: the system is in one direction a superfluid and in the other still a solid. To complete this outline, when the coupling constant is further increased there is yet another quantum phase transition associated with the proliferation of disclinations turning the system into an isotropic superfluid.

Let us now review the "dislocation duality": in close analogy with vortex duality, this shows how crystals and liquid crystals are related via a weak–strong duality. The requirement that disclinations have to be kept out of the vacuum is actually a greatly simplifying factor. One follows the same dualization procedure for the dislocations as for the vortices. Hence, we introduce Hubbard–Stratonovich auxiliary tensor fields $\sigma_{\mu\nu}$, rewriting the action as,

$$S = \int \mathrm{d}\tau \mathrm{d}x^2 \left[\frac{1}{4\mu} \left(\sigma_{\mu\nu}^2 - \frac{\nu}{1+\nu} \sigma_{\mu\mu}^2 \right) + \mathrm{i}\sigma_{\mu\nu} w_{\mu\nu} \right] , \qquad (6.11)$$

where $\nu = \lambda/2(\lambda+\mu)$ is the Poisson ratio. We divide the displacement fields (having the same status as the phase field in vortex duality) in smooth and multivalued parts $u_\mu = u_\mu^{\text{smooth}} + u_\mu^{\text{MV}}$, and integrating out the smooth strains yields a constraint, in this case a Bianchi identity,

$$\partial_\mu \sigma_{\mu\nu} = 0 , \qquad (6.12)$$

The physical meaning of $\sigma_{\mu\nu}$ is that they are the stress fields, which are conserved in the absence of external stresses as in Eq. (6.12): the above is just the stress–strain duality of elasticity theory. One now wants to parametrize the stress fields in terms of a gauge field. Since the stress tensor is symmetric this is most naturally accomplished in terms of Kleinert's double curl gauge fields,

$$\sigma_{\mu\nu} = \epsilon_{\mu\kappa\lambda}\epsilon_{\nu\kappa'\lambda'}\partial_\kappa\partial_{\kappa'}B_{\lambda\lambda'} \tag{6.13}$$

while the B's are *symmetric* tensors, otherwise transforming as $U(1)$ gauge fields.

To maintain the analogy with the vortex duality as tightly as possible, one can as well parametrize it in a normal gauge field, $\sigma_{\mu\nu} = \varepsilon_{\mu\kappa\lambda}\partial_\kappa b^\nu_\lambda$ with the requirement that one has to impose the symmetry of the stress tensor explicitly by Lagrange multipliers. Using this route one finds that the multivalued strains turn into a source term $ib^\nu_\mu J^\nu_\mu$ where,

$$J^{\mathrm{V}}_{\mu\nu} = \epsilon_{\mu\kappa\lambda}\partial_\kappa\partial_\lambda u^{\mathrm{MV}}_\nu , \tag{6.14}$$

This is just like a vortex current carrying an extra "flavour" ν. It is the dislocation current, where the flavour indicates the $D+1$ components of the Burgers vector. Like the vortices, dislocations have long-range interactions which are parametrized by the gauge fields b (or B), with the special effect that these are only active in the directions of the Burgers vectors.

The double curl gauge fields have the advantage that the symmetry is automatically built in while the "extra derivatives" enable the identification of the disclination currents. One finds,

$$S = \int \mathrm{d}\tau \mathrm{d}x^2 \left[\frac{1}{4\mu}\left(\sigma^2_{\mu\nu} - \frac{\nu}{1+\nu}\sigma^2_{\mu\mu}\right) + \mathrm{i}B_{\mu\nu}\eta_{\mu\nu} \right] , \tag{6.15}$$

where the "stress gauge fields" B are sourced by a total "defect current",

$$\begin{aligned}\eta_{\mu\nu} &= \epsilon_{\mu\kappa\lambda}\epsilon_{\nu\kappa'\lambda'}\partial_\kappa\partial_{\kappa'}w^{\mathrm{MV}}_{\lambda\lambda'} , \\ &= \theta_{\mu\nu} - \epsilon_{\mu\kappa\lambda}\partial_\kappa J_{\nu\lambda} ,\end{aligned} \tag{6.16}$$

where $\theta_{\mu\nu}$ is the disclination current, and ν refers to the Franck vector component. The fact that the disclination current has "one derivative less" than the dislocation current actually implies that disclinations are in the solid confined—in the solid, a disclination is like a quark.

One now associates a much larger core energy to the disclinations than to the dislocations, and upon increasing the coupling constant a loop blowout transition will occur involving only the dislocation world lines—it is obvious from the single curl gauge field formulation that dislocations are just like vortices carrying an extra "Burgers flavour". To obtain the quantum nematic one populates all Burgers directions equally and after some straightforward algebra one obtains the effective action for the "Higgsed stress photons" having the same status as Eq. (3.23) for the Mott insulator,

$$S = \int d\tau dx^2 \left[m_{\mathrm{nem}}^2 \sigma_{\mu\nu} \frac{1}{\partial^2} \sigma_{\mu\nu} + \frac{1}{4\mu} \left(\sigma_{\mu\nu}^2 - \frac{\nu}{1+\nu} \sigma_{\mu\mu}^2 \right) + \mathrm{i} B_{\mu\nu} \theta_{\mu\nu} \right] , \tag{6.17}$$

where σ should be expressed in the double curl gauge field $B_{\mu\nu}$ according to Eq. (6.13). In terms of the regular gauge fields b_{μ}^{ν}, the first term represents a Higgs mass, while the second term is like a Maxwell term. Nevertheless, in the nematic the disclinations still act as sources coupling to the double curl gauge fields.

Ignoring the disclinations, one finds in 2+1d that Eq. (6.17) describes a state is quite similar to a Mott insulator: all excitations are massive, and one finds now a triplet of massive "photons". These are counted as follows: there are two propagating (longitudinal and transversal acoustic) phonons of the background world crystal, turning into "stress photons" after dualization and acquiring a mass in the nematic. In addition, the dislocation condensate adds one longitudinal stress photon.

As it turns out, the rules change drastically upon breaking the Lorentz invariance. In a crystal formed from material bosons, displacements in the time direction u_τ are absent, and this has among others the consequence that the dislocation condensate does not couple to compressional stress. Instead of the incompressible nature of the relativistic state, one finds now two massless modes in the quantum nematic: a rotational Goldstone boson associated with the restoration of the broken rotational symmetry, and a massless sound mode which can be shown to be just the zero sound mode of the superfluid. The non-relativistic quantum liquid crystals are automatically superfluids as well and their relation to gravity is obscured.

Turning to the 3+1d case one finds as extra complication that dislocations turn into strings and one has to address the fact that the "stress superconductor" is now associated with a condensate of strings. One meets the same complication as in vortex duality, which was tackled in chapter 3. The outcome is actually quite straightforward: the effective London actions of the type Eqs. (3.23),(6.17) have the same form regardless whether one deals with particle or string condensates, and these enter through the Higgs term $\sim \sigma^2/\partial^2$.

How to interpret the 2+1d relativistic quantum nematic? There are no low energy excitations and it only reacts to disclinations. It has actually precisely the same status as a flat Einsteinian spacetime in 2+1d that only feels the infinitesimal vibrations associated with gravitational events far away. Similarly, using the general relativity (GR) technology of the next section, it is also straightforward to demonstrate [102, 103] that in 3+1d one ends up with two massless spin-2 modes: the gravitons. To prove that it is precisely linearized gravity, let us consider next the rules of Kleinert that allow to explicitly relate these matters to gravitational physics.

6.2.2 Quantum elasticity field theory: the Kleinert rules

Elaborating on a old tradition in "mathematical metallurgy", Kleinert identified an intriguing portfolio of general correspondences between the field theory describing elastic media and the geometrical notions underlying general relativity. In order to appreciate what comes, we need to familiarize the reader with some of the entries of this dictionary. For an exhaustive exposition, see Kleinert's books on the subject [41, 42].

GR is a geometrical theory which departs from a metric $g_{\mu\nu}$, such that an infinitesimal distance is measured through,

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x_\mu \mathrm{d}x_\nu \,, \tag{6.18}$$

One now insists that the physics is invariant under local coordinate transformations (general covariance) $x_\mu \to \xi_\mu(x_\nu)$; infinitesimal transformations then are like gauge transformations of the metric,

$$g_{\mu\nu} \to g_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu \equiv g_{\mu\nu} + h_{\mu\nu} \,, \tag{6.19}$$

Only quantities are allowed in the theory which are invariant under these transformations and insisting on the minimal number of gradients, one is led to the Einstein–Hilbert action governing spacetime,

$$S = -\frac{1}{2\kappa}\int \mathrm{d}^D x \mathrm{d}t \; R\sqrt{-g} \,, \tag{6.20}$$

where $g = \det g_{\mu\nu}$ and R the Ricci scalar, while κ is set by Newton's constant. Together with the part describing the matter fields, the Einstein equations follow from the saddle points of this action.

How to relate this to solids? Imagine that one lives inside a solid and all one can do to measure distances is to keep track how one jumps from unit cell to unit cell. In this way one can define a metric "internal" to the solid, and the interesting question becomes: what is the fate of the diffeomorphisms ("diffs") Eq. (6.19)? In order to change the metric one has to displace the atoms and this means that one has to *strain* the crystal,

$$g_{\mu\nu} \rightarrow g_{\mu\nu} + w_{\mu\nu} \,, \tag{6.21}$$

But the strain fields are surely not gauge fields: the elastic energy Eq. (6.9) explicitly depends on the strain. Obviously, the crystal is non-diffeomorphic and it is characterized by a "preferred" or "fixed" frame. This is the deep reason that normal crystals have nothing to do with GR.

In standard GR the objects that are invariant keep track of curvature and these appear in the form of curvature tensors in the Einstein equations. Linearizing these, assuming only infinitesimal diffs as in Eq. (6.19), one finds for the Einstein tensor appearing in the Einstein equations, say in the 2+1d case to avoid superfluous labels,

$$G_{\mu\nu} = \epsilon_{\mu\kappa\lambda}\epsilon_{\nu\kappa'\lambda'}\partial_\kappa\partial_{\kappa'}h_{\lambda\lambda'} \tag{6.22}$$

One compares this with the disclination current Eq. (6.16) and one discovers that these are the same expressions after associating the strains $w_{\mu\nu}$ with the infinitesimal diffs $h_{\mu\nu}$. This is actually no wonder: at stake is that the property of curvature is independent of the gauge choice for the metric. One can visualize the curved manifold in a particular gauge fix, and this is equivalent to the fixed frame. The issue is that curvature continues to exist when one lets loose the metric in the gauge volume.

What is the meaning of the dislocation tensor? Cartan pointed out to Einstein that his theory was geometrically incomplete: one has to allow also for the property of torsion. It turns out that torsion is "Cartan-Einstein" GR sourced by spin currents and the effects of it turn out to be too weak to be observed (see e.g. Ref. [115]). In the present context, the torsion tensor appearing in the equations of motions precisely corresponds with the dislocation currents. With regard to these topological aspects, crystals and spacetime are remarkably similar.

However, given the lack of general covariance the dynamical properties of spacetime and crystals are entirely different. For obvious reasons, spacetime does not know about phonons while crystals do not know about gravitons, let alone about black holes. A way to understand why things go so wrong is to realize that the disclinations encode for curvature, and gravitons can be viewed as infinitesimal curvature fluctuations. As we already explained, disclinations are confined in crystals meaning that it costs infinite energy to create curvature fluctuations in normal solids.

Let us now turn to the relativistic quantum nematics: here the situation looks much better. Gravity in 2+1d is incompressible in the sense that the constraints do not permit massless propagating modes, the gravitons. We also found out that disclinations are now deconfined and they appear as sources in the effective action Eq. (6.17): this substance knows about curvature. In fact, one can apply similar considerations to the 3+1d case, where two massless spin-2 modes are present. The relativistic quantum nematic in 3+1d behaves quite like spacetime!

To make the identification even more precise, one notices that the expression for the linearized Einstein tensor Eq. (6.22) is coincident with the expression for the stress tensor in terms of the double curl gauge field $B_{\mu\nu}$, Eq. (6.13). But now one is dealing with gauge invariance both of $B_{\mu\nu}$ and $h_{\mu\nu}$ while they are both symmetric tensors. At least on the linearized level the stress tensor *is* the Einstein tensor. It is now easy to show that the Higgs term in the theory of the nematic when expressed in terms of the linearized Einstein tensors,

$$\begin{aligned}\sigma_{\mu\nu}\frac{1}{\partial^2}\sigma_{\mu\nu} &= G_{\mu\nu}\frac{1}{\partial^2}G_{\mu\nu}\\ &\to R,\end{aligned} \tag{6.23}$$

actually reduces to the Ricci tensor R, demonstrating that one recovers the Einstein–Hilbert action at distances large compared to the Higgs scale. Once more, this only holds in the linearized theory. This works in the same way in 3+1 (and higher) dimensions which is the easy way to demonstrate that gravitons have to be present [102, 103]. At least the linearized version of the Einstein–Hilbert action appears to be precisely coincident with the effective field theory describing the collective behaviour of the quantum nematic!

Although this all looks convincing there is still a gap in the conceptual understanding of what has happened with the geometry of the crystal in the presence of condensed dislocations. The emergence of gravity requires that the original spacetime defined by the crystal has to become diffeomorphic.

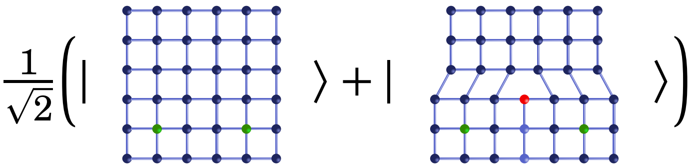


Figure 6.4: In the dislocation condensate (quantum nematic), the distance between two points (green dots) is in a coherent superposition of having zero, one or any number of half-line insertions (light blue) or dislocations (red dot) in between them, and therefore the number of lattice spacings in between them is undefined. This is equivalent to having the Einstein translations fully gauged: there is a diffeomorphism between configurations with any number of lattice spacings in between the two points.

The fields as of relevance to the dynamics of the nematic are healthy in this regard but they belong to the dual side. The analogy with the Mott insulator is now helpful: to demonstrate that gravity has emerged requires the demonstration that the spacetime of the original crystal is diffeomorphic and that is equivalent to demonstrating that in the vortex condensate the superfluid phase acquires a compact $U(1)$ gauge invariance. The diffeomorphic nature of the stress gauge fields telling about the excitations of the quantum nematic has in turn the same status as the gauge fields that render the vortex condensate to be a superconductor.

The good news is that we can use the same "first quantization" trick that helped us to understand the emergence of the stay-at-home gauge in the vortex condensate to close this conceptual gap. As for the vortices, it is easy to picture what happens to the metric of the crystal when the coherent superpositions of dislocation configurations associated with the dual stress superconductor are present. Let us repeat the exercise at the end of the previous section (Fig. 6.1), by comparing how two points some distance apart communicate with each other, but now focusing on the metric properties. This is illustrated in Fig. 6.4: imagine that no dislocation is present between the two points and one needs N jumps to get from one point to the other. However, this configuration is at energies less than the Higgs mass of the quantum nematic necessarily in coherent superposition with a configuration where a dislocation has moved through the line connecting the two points:

one now needs $N+1$ hops and since these configurations are in coherent superposition "$N = N+1$" and the geometry is now truly diffeomorphic!

However, there is one last caveat. Although translational symmetry is restored in the quantum nematic, the rotations are still in a fixed frame and even spontaneously broken! This is different from full Einstein gravity: in real spacetime also the Lorentz transformations (rotations in our Euclidean setting) are fully gauged. In order to understand this point, let us start from special relativity, which has the global symmetry of the Poincaré group comprising translations and Lorentz transformations. The translations form a subgroup, such that translational and rotational symmetry are easily distinguishable. More precisely, the generators of translations are ordinary derivatives ∂_μ which commute $[\partial_\mu, \partial_\nu] = 0$. In many ways, going from special to general relativity is from going from global to local Poincaré symmetry [115]. Indeed, referring to elasticity language, it seems to make sense to restore first translational and then rotational symmetry, ending up in a perfectly locally symmetric "liquid" state.

However, it has long been known that such "gauging of spacetime symmetry" is very intricate, which has to do with the definition of locality under such transformations. What happens is that local coordinate transformations of the form $x_\mu \to \xi_\mu(x_\nu)$, which are in fact local translations, can also correspond to local rotations. The local translations no longer form a subgroup, as the generators of translations should be augmented to those of *parallel translations*, defined by [116],

$$D_\mu = \partial_\mu + \Gamma_\mu^{\ \kappa\lambda} f_{\kappa\lambda}, \tag{6.24}$$

where $\Gamma_\mu^{\ \kappa\lambda}$ is the connection and $f_{\kappa\lambda}$ is the generator of local rotations. Such modified derivatives do not commute, and two consecutive translations may result in a finite rotation. Such symmetry structure is actually at the heart at everything non-linear happening in Einstein theory including black holes.

Going back to what we now know of the quantum nematic, it is clear that it cannot correspond to full GR, since rotational symmetry as reflected by disclinations is still gapped. Nevertheless, the identification between quantum nematics and *linearized* gravity is in perfect shape. Linearized gravity is a special and somewhat pathological limit of full GR, as it only applies to nearly globally Lorentz symmetric systems. It was quite some time ago realized that such systems are symmetric under global Lorentz transformations and infinitesimal coordinate transformations (see ch. 18,35 in Ref. [117]).

This is equivalent to fixing the Lorentz frame globally yet allow for infinitesimal Einstein translations. Under such conditions the equations of motion of linearized gravity follow automatically.

Here we have demonstrated that linearized gravity—a very peculiar limiting case of GR—is actually literally realized in a quantum nematic. The deeper reason is that in a quantum nematic the rotational symmetry of (Euclidean) spacetime is global and even spontaneously broken, while the restoration of the translational symmetry by the dislocation condensate has caused the fixed frame internal coordinate system of the crystal to turn into a geometry that is characterized by a covariance exclusively associated with infinitesimal translational coordinate transformations.

6.3 Summary and outlook

In so far as vortex duality is concerned we have presented here no more than a clarification. Living on the "dual side", where the Bose-Mott insulator appears as just a relativistic superconductor formed from vortices, the emerging stay-at-home local charge conservation from the canonical representation in terms of the Mott insulating phase of the Bose-Hubbard model is not manifestly recognizable. However, the dual vortex language contains all the information required to reconstruct precisely the nature of the field configurations of the "original" superfluid phase fields which are realized in the vortex superconductor. By inspecting these we identified a very simple but intriguing principle. The local charge conservation of the Mott insulator, associated with the emergent stay-at-home compact $U(1)$ gauge symmetry, is generated in the vortex condensate by the quantum mechanical principle that states in coherent superposition "are equally true at the same time"—the Schrödinger cat motive.

We find this simple insight useful since it yields a somewhat more general view on the nature of strong/weak dualities. We already emphasized that Mott insulators as defined through the local conservation of charge do not necessarily need a lattice. One does not have to dig deep to find an example: our dual superconductor is just a relativistic superconductor in 2+1d, which is in turn dual to a Coulomb phase that can also be seen as a superfluid. The charge associated with this superfluid is locally conserved in the superconductor, regardless of whether the superconductor lives on a lattice or in the continuum.

We find the emergent gauging of translational symmetry realized in the quantum nematic an even better example of the usefulness of this insight. Earlier work indicated that the relativistic version of this nematic is somehow associated with emergent gravity. Resting on the "coherent superposition" argument it becomes directly transparent what causes the gauging of the crystal coordinates: the condensed dislocations "shake the coordinates coherently" such that infinitesimal Einstein translations appear while the Lorentz frame stays fixed. This emergent symmetry imposes that the collective excitations of the quantum nematic have to be in one-to-one correspondence with linearized gravity.

Our message is that we have identified a mechanism for the "dynamical generation" of gauge symmetry which is very simple but also intriguing viewed from a general physics perspective: the quantum mechanics principle of states in coherent superposition being "equally true at the same time" translates directly to the principle that the global symmetry that is broken in the ordered state is turned into a gauge symmetry on the disordered side just by the quantum undeterminedness of the topological excitations in the dual condensate. This raises the interesting question: is quantum coherence required for the emergence of local symmetry, or can it also occur in classical systems?

This question relates directly to the spectacular recent discovery of "Dirac monopoles" in spin ice [118]. Castelnovo, Moessner and Sondhi [119] realized that the manifold of ground states ("frustration volume") of this classical geometrically frustrated spin problem is coincident with the gauge volume of a compact $U(1)$ gauge theory, with the ramification that it carries Dirac monopoles as topological excitations. All along it has been subject of debate to what extent these monopoles can be viewed as literal Dirac monopoles in the special "vacuum" realized in the spin ice, or rather half-bred cartoon versions of the real thing. With our recipe at hand it is obvious how to make them completely real: imagine the classical spin ice to fill up Euclidean spacetime, and after Wick rotation our "coherent superposition principle" would have turned the frustration volume of the classical problem into a genuine gauge volume since by quantum superposition all degenerate states would be "equally true at the same time".

The ambiguity associated with the classical spin ice monopoles is rooted in the role of time. In principle, by doing time-resolved measurements one can observe every particular state in the frustration volume and this renders these states to be not gauge equivalent. However, all experiments which have revealed the monopoles involved large, macroscopic time scales. One can pose the question whether it is actually possible under these conditions to define observables that can discriminate between the “fake” monopoles of spin ice and the monopoles of Dirac. Perhaps the answer is pragmatic: as long as ergodicity is in charge, one can rely on the ensemble average instead of the time average, and as long as the time scale of the experiment is long enough such that one is in the ergodic regime, the frustration volume will “disappear” in the ensemble average. For all practical purposes one is then dealing with a genuinely emergent gauge symmetry which tells us that in every regard the spin ice monopole is indistinguishable from the Dirac monopole.