Multicritical Symmetry Breaking and Naturalness of Slow Nambu-Goldstone Bosons
Griffin, Tom; Grosvenor, Kevin John Torres; Horava, Petr; Yan, Ziqi

Published in:
Physical Review D (Particles, Fields, Gravitation and Cosmology)

DOI:
10.1103/PhysRevD.88.101701

Publication date:
2013

Citation for published version (APA):
Gapless Nambu-Goldstone (NG) modes \[1–4\] appear prominently across an impressive array of physical phenomena, both relativistic and nonrelativistic (for reviews, see e.g. [5,9]). They are a robust consequence of spontaneous symmetry breaking. Moreover, when further combined with gauge symmetries, they lead to the Higgs phenomenon, responsible for controlling the origin of elementary particle masses.

The NG modes are controlled by Goldstone’s theorem: A spontaneously broken generator of a continuous internal rigid symmetry implies the existence of a gapless mode. With Lorentz invariance, the theorem implies a one-to-one correspondence between the generators of broken symmetry and massless NG modes, but in the nonrelativistic setting, it leaves questions [10,12]: What is the number of independent NG modes? What are their low-energy dispersion relations?

In this paper, we study the general classification of NG modes, and their Naturalness, in nonrelativistic theories with Lifshitz symmetries. The important concept of Naturalness is behind many successes of modern physics, but it also leads to some of its most intriguing and persistent puzzles. A system is technically natural if its low-energy behavior follows from that at higher energy scales, without requiring fine tuning [13]. Perhaps the most famous “Naturalness problem” comes from the apparent smallness of the cosmological constant [14,15], suggesting that something fundamental is still missing in our understanding of gravity and cosmology. And now that the Higgs boson has been discovered, (un)naturalness at the TeV scale is again at the forefront of high-energy particle physics [17,21]. In the context of quantum gravity, theories with Lifshitz symmetries have been studied at least in part because of their improved short-distance (UV) behavior [22,24]. Our study illustrates that in Lifshitz-type theories, not only the short-distance behavior but also the concept of Naturalness acquires interesting new features.

Effective field theory and Goldstone’s theorem: In [25,26], elegant arguments based on effective field theory (EFT) have been used to clarify the consequences of Goldstone’s theorem in the absence of Lorentz invariance. The main idea is to classify possible NG modes by classifying the EFTs available for describing their low-energy dynamics. We start with the NG field components \(\pi^A\), \(A = 1, \ldots, n\), which serve as coordinates on the space of possible vacua \(M = \mathcal{G}/H\) in a system with symmetries broken spontaneously from \(\mathcal{G}\) to \(H \subset \mathcal{G}\). Our spacetime will be the flat \(\mathbb{R}^{D+z}\) with coordinates \(t, x^i, i = 1, \ldots, D\), and we impose the Lifshitz symmetry consisting of all Euclidean isometries of the spatial \(\mathbb{R}^D\) and the time translations. At the fixed points of the renormalization group, this symmetry is enhanced by anisotropic scaling symmetry \(x^i \to bx^i, t \to b^{2/z} t\), with the dynamical exponent \(z\) characterizing the degree of anisotropy at the fixed point.

Arguments of [25,26] suggest that the generic low-energy EFT action for the NG fields \(\pi^A\) with these symmetries is

\[
S_{\text{eff}} = \frac{1}{2} \int dt d^D x \left\{ \Omega_A(\pi) \dot{\pi}^A + g_{AB}(\pi) \dot{\pi}^A \dot{\pi}^B - \eta_{AB} \right\}, \tag{1}
\]

where \(\Omega_A\), \(g_{AB}\) and \(\eta_{AB}\) are backgrounds transforming appropriately under \(\mathcal{G}\) and “…” stands for higher-order derivative terms. The term linear in \(\dot{\pi}^A\) is only possible because of the special role of time. Lorentz invariance would require \(\Omega_A = 0\) and \(g_{AB} = h_{AB}\), thus reproducing the standard relativistic result: One massless, linearly dispersing NG mode per each broken symmetry generator. In the nonrelativistic case, turning on \(\Omega_A\) leads to \(t^2\) types of NG bosons [25,26]: First, those field components that get their canonical momentum from \(\Omega_A\) form canonical pairs; each pair corresponds to a pair of broken generators, and gives one Type-B NG mode with a quadratic dispersion. The remaining, Type-A modes then get their canonical momenta from the second term in [1], and behave as in the relativistic case, with \(z = 1\). In both cases, higher values of \(z\) can arise if \(h_{AB}\) becomes accidentally degenerate [25].

We will show that in Lifshitz-type theories, \(h_{AB}\) can be small naturally, without fine tuning. When that happens, the low-energy behavior of the NG modes will be
determined by the next term, of higher order in \( \partial_t \). The argument can be iterated: When the terms of order \( \partial^4 \) are also small, terms with \( z = 3 \) will step in, etc. This results in a hierarchy of multicritical Type-A and Type-B NG modes with increasing values of \( z \). Compared to the generic NG modes described by \( \begin{bmatrix} 1 \end{bmatrix} \), these multicritical NG modes are anomalously slow at low energies.

\( z = 2 \) linear and nonlinear \( O(N) \) sigma models:

We will demonstrate our results by focusing on a simple but representative example of symmetry breaking, the \( O(N) \) nonlinear sigma model (NLSM) with target space \( S^{N-1} \). (For some background on Lifshitz scalar theories, see \[22, 27–31\].) Until stated otherwise, we will also impose time reversal invariance, to forbid \( \Omega \). The action of the \( O(N) \)-invariant \( z = 2 \) Lifshitz NLSM \[29\] is then

\[
S_{\text{NLSM}} = \frac{1}{2G^2} \int dt d^D x \left\{ g_{AB} \pi^A \pi^B - c^2 g_{AB} \Delta \pi^A \Delta \pi^B - \lambda_1 (\pi^A \partial \pi^B) (\pi^C \partial \pi^D) - \lambda_2 (g_{AB} \partial \pi^A \partial \pi^B)^2 - c^2 g_{AB} \partial \pi^A \partial \pi^B \right\}. \tag{2}
\]

Here \( \Delta \pi^A = \partial \pi^A + \Gamma_{A}^{BC} \partial \pi^B \partial \pi^C \), \( g_{AB} \) is the round metric on the unit \( S^{N-1} \) (later we will use \( g_{AB} = \delta_{AB} + \pi^A \pi^B / (1 - \delta_{CD} \pi^C \pi^D) \)), and \( \Gamma_{A}^{BC} \) is its connection. The Gaussian \( z = 2 \) RG fixed point is defined by the first two terms in \( \{2\} \) as \( G \rightarrow 0 \). We define scaling dimensions throughout in the units of spatial momentum, \( [\partial] = 1 \). Due to its geometric origin, the NG field \( \pi^A \) is dimensionless, \( [\pi^A] = 0 \). The first four terms in \( S_{\text{NLSM}} \) are all of the same dimension, so \( [c^2] = [\lambda_1] = [\lambda_2] = 0 \). We can set \( c = 1 \) by the rescaling of space and time, and will do so throughout the paper. All interactions are controlled by the coupling constant \( G \), whose dimension is \( [G] = (2 - D)/2 \). Thus, the critical spacetime dimension of the system, at which the first four terms in \( \{2\} \) are classically marginal, is equal to \( 2 + 1 \). The remaining term has a coupling of dimension \( [c^2] = 2 \), and represents a relevant deformation away from \( z = 2 \), even in the noninteracting limit \( G \rightarrow 0 \). Since \( c \) determines the speed of the NG modes in the \( k \rightarrow 0 \) limit, we refer to this term as the “speed term” for short. Given the symmetries, this relevant deformation is unique.

We are mainly interested in \( 3 + 1 \) dimensions, so we set \( D = 3 \) from now on. Since this is above the critical dimension of \( 2 + 1 \) and \( [G] \) is negative, the theory described by \( \{2\} \) must be viewed as an EFT: \( S_{\text{NLSM}} \) gives the first few (most relevant) terms out of an infinite sequence of operators of growing dimension, compatible with all the symmetries. It is best to think of this EFT as descending from some UV completion. For example, we can engineer this effective NLSM by starting with the \( z = 2 \) linear sigma model (LSM) of the unconstrained \( O(N) \) vector \( \phi^I \), \( I = 1, \ldots, N \), and action

\[
S_{\text{LSM}} = \frac{1}{2} \int dt d^3 x \left\{ \phi^I \phi^I - c^2 \partial^2 \phi^I \partial^2 \phi^I - c^2 \partial \phi^I \partial \phi^I \right\}
\]

The first two terms define the Gaussian \( z = 2 \) fixed point. For example, we can engineer \( \Delta \pi^A = \partial \pi^A + \Gamma_{A}^{BC} \partial \pi^B \partial \pi^C \), who's dimension is \( [\pi^A] = 2 \) by rescaling space and time. At this fixed point, the field is of dimension \( [\phi] = 1/2 \), and the dimensions of the couplings – in the order from the marginal to the more relevant – are: \( [c] = [g_3] = [g_2] = [f_2] = 0 \), \( [g_4] = [f_1] = 1 \), \( [g_2] = [c^2] = 2 \), \( [\lambda] = 3 \) and \( [m^4] = 4 \).

\[2\]

\[3\]

\[4\]

\[5\]
of divergence of a diagram with $L$ loops, $E$ external legs and $V_3$ cubic vertices is $D = 8 - 2E - 3L - 2V_3$. Loop corrections to the speed term are actually finite. If we start at the classical level by setting $c^2 = 0$, this relation can be viewed as a “zeroth order natural relation” (in the sense of [32]): True classically and acquiring only finite corrections at all loops. We can even set $c^2$ at any order to zero by a finite local counterterm, but an infinite counterterm for $c^2$ is not needed for renormalizability.

How large is this finite correction to $c^2$? At one loop, five diagrams (shown in Fig. 1) contribute to the inverse propagator $\Gamma_{AB}(\omega, k) \equiv (\omega^2 - |k|^2 + \Sigma(\omega, k))\delta_{AB}$. We can read off the one-loop correction to $c^2 = 0$ by expanding $\Sigma = -\delta m^4 - \delta c^2 k^2 + \ldots$. Four of these diagrams give a (linearly) divergent contribution to $\delta m^4$, but both the divergent and finite contributions to $\delta m^4$ sum to zero, as they must by Goldstone’s theorem. The next term in $\Sigma$ is then proportional to $k^2$ and finite. It gets its only one-loop contribution from diagram (d) in Fig. 1 whose explicit evaluation gives

$$\delta c^2 = \frac{27/4 \cdot 5}{63\pi^5/2} \left[ \Gamma \left( \frac{1}{4} \right) \right]^2 \frac{\lambda}{m} \approx 0.0125 \frac{\lambda}{m}. \quad (6)$$

Thus, the first quantum correction to $c^2$ is indeed finite and nonzero. But is it small or large? There are much higher scales in the theory, such as $\Lambda$ and $\epsilon$, yet in our weak coupling limit the correction to the speed term is found to be $\delta c^2 \propto \lambda / m$ naturally. In this sense, $\delta c^2$ is small, and so $c^2$ can also be small without fine tuning.

We can also calculate $\delta c^2$ at one loop in the effective NLSM. The Feynman rules derived from [2] for the rescaled field $\pi^A/G$ involve a propagator independent of $G$ (in which we set $c^2 = 0$), and an infinite sequence of vertices with an arbitrary even number of legs, of which we will only need the lowest one. When the radial direction of $\phi$ is integrated out in our superrenormalizable LSWM, at the leading order we get [2] with $G = 1/v$, $\lambda_1 = 0$ and $\lambda_2 = 1$. In this special case, the 4-vertex is

$$D_{\beta \alpha} k_0 k_1 k_2 k_3 \left( \omega_1 + \omega_2 \omega_3 + \omega_4 \right) + (k_1 + k_2)^2 (k_3 + k_4)^2 \right) \delta_{AB} \delta_{CD} + 2 \text{ permutations.}$$

The first quantum correction to $\delta c^2$ comes at one loop, from $\bigcirc \bigcirc$, and it is cubically divergent. With the sharp cutoff at $|k| = \Lambda$, we get

$$\delta c^2 = \frac{G^2 \Lambda^3}{3\pi^2}. \quad (7)$$

This theory is only an EFT, and its natural cutoff scale $\Lambda$ is given by $m$, the gap scale of the $\sigma$. With this value of the cutoff, the one-loop result (7) gives $\delta c^2 = O(\lambda / m)$, which matches our LSWM result.

If one wishes to extend the control over the LSW beyond weak coupling in $\lambda$, one can take the large-$N$ limit, holding the ’t Hooft coupling $\lambda N$ fixed. In this limit, the LSWM and the NLSM actually become equivalent, by the same argument as in the relativistic case [33]. An explicit calculation shows that at large $N$, $\delta c^2$ is not just finite but actually zero, to all orders in the ’t Hooft coupling.

**Naturalness:** Now, we return to the question of Naturalness of small $\delta c^2$, in the technical context articulated in [13]. As a warm-up, consider first our superrenormalizable LSWM in its unbroken phase. The leading contribution to the speed term in the inverse propagator of $\phi'$ is now at two-loop order, from $\bigotimes$. This diagram is finite; even the leading constant, independent of $\omega$ and $k$, only yields a finite correction to the gap $m^4$. The contribution of order $k^2$ is then also finite, and gives $\delta c^2 = \xi \lambda^2 / m^4$, with $\xi$ a pure number independent of all couplings. But is this $\delta c^2$ small?

Let us first recall a well-known fact from the relativistic $\lambda \phi^4$ theory [13]: $\lambda$ and $m^2$ may be simultaneously small, $\sim \epsilon$, because in the limit of $\epsilon \to 0$, the system acquires an enhanced symmetry – in this case, the constant shift symmetry,

$$\phi' \to \phi' + \alpha' \quad (8)$$

The same constant shift symmetry works also in our superrenormalizable Lifshitz LSWM. Restoring dimensions, we have

$$\lambda = O(\epsilon \mu^3), \quad \quad m^4 = O(\epsilon \mu^4). \quad (9)$$

Here $\mu$ is the scale at which the constant shift symmetry is broken (or other new physics steps in), and represents the scale of naturalness: The theory is natural until we reach the scale $\mu = O(m^4 / \lambda)$. This result is sensible – if we wish for the scale of naturalness to be much larger than the gap scale, $\mu \gg m$, we must keep the theory at weak coupling, $\lambda / m^3 \ll 1$. Now, how about the speed term? If we assume that $\delta c^2$ is also technically small, $c^2 \sim \epsilon$, this assumption predicts $\delta c^2 = O(\lambda^2 / m^4)$, which is exactly the result we found above in our explicit perturbative calculation. It looks like there must be a symmetry at play, protecting simultaneously the smallness of $m^4$, $\lambda$ as well as $c^2$! We propose that the symmetry in question is the generalized shift symmetry [8], with $a'$ now a quadratic polynomial in the spatial coordinates,

$$a' = a'_x x^4 + a'_y x^4 + a'_4. \quad (10)$$
The speed term \( \partial_i \phi^I \partial_i \phi^I \) is forbidden by this “quadratic shift” symmetry, while \( \partial^2 \phi^I \partial^2 \phi^I \) is invariant up to a total derivative. This symmetry holds in the free-field limit, and will be broken by interactions. It can be viewed as a generalization of the Galileon symmetry, much studied in cosmology [34], which acts by shifts linear in the spacetime coordinates.

As long as the coupling is weak, the unbroken phase of the LSM exhibits a natural hierarchy of scales, \( c \ll m \ll \mu \), with the speed term much smaller than the gap scale. The effects of the speed term on the value of \( z \) would only become significant at low-enough energies, where the system is already gapped. Note that another interesting option is also available, since there is no obligation to keep \( c \) small at the classical level. If instead we choose \( c \) much above the gap scale \( m \) (but below the naturalness scale \( \mu \)), we can go to lower energies the system will experience a crossover from \( z = 2 \) to \( z = 1 \) before reaching the gap, and the theory will flow to the relativistic \( \lambda \phi^4 \) in the infrared. The coupling \( \lambda \) can stay small throughout the RG flow from the free \( z = 2 \) fixed point in the UV to the \( z = 1 \) theory in the infrared.

Now consider the same LSM in the broken phase. In this case, we are not trying to make \( m \) small – this is a fixed scale, setting the nonzero gap of the \( \sigma \). Moreover, the \( \pi \)'s are gapless, by Goldstone’s theorem. We claim that \( c^2 \) can be naturally small in the regime of small \( c \):

\[
\lambda = O(\varepsilon^3), \quad c^2 = O(\varepsilon \mu^2),
\]

as a consequence of an enhanced symmetry. The symmetry in question is again the “quadratic shift” symmetry, now acting only on the gapless \( \sigma \) modes in their free-field limit: \( \Pi^A \to \Pi^A + a^A_{ij} x^i x^j + \ldots \). It follows from \([1]\) that the radius \( v \) of the vacuum manifold \( S^{N-1} \) goes to infinity with \( \varepsilon \to 0 \), \( v = O(m^2/\sqrt{\mu^2} \varepsilon) \), and \( v \to \infty \) corresponds to the free-field limit of the \( \pi \)'s. Our enhanced symmetry does not protect \( m \) from acquiring large corrections; we can view \( m \) in principle as a separate mass scale, but it is natural to take it to be of the order of the naturalness scale, \( m = O(\mu) \). Altogether, this predicts \( \delta c^2 = O(\lambda / \mu) = O(\lambda / m) \), in accord with our explicit loop result \([6]\).

The technically natural smallness of the speed term in our examples is not an artifact of the superrenormalizability of our LSM. To see that, consider the full renormalizable LSM \([9]\), first in the unbroken phase. As we turn off all self-interactions by sending \( \varepsilon \to 0 \), the enhanced quadratic shift symmetry will again protect the smallness of \( c^2 \sim \varepsilon \). In terms of the naturalness scale \( \mu \), this argument predicts that in the action \([4]\), all the deviations from the \( z = 2 \) Gaussian fixed point can be naturally of order \( \varepsilon \) in the units set by \( \mu \):

\[
e_2 = O(\varepsilon), \quad \ldots, \quad c^2 = O(\varepsilon \mu^2), \quad \lambda = O(\varepsilon \mu^3), \quad m^4 = O(\varepsilon \mu^4).
\]

If we want the naturalness scale to be much larger than the gap scale, \( \mu \gg m \), all couplings must be small; for example, \( e_2 = O(m^4/\mu^4) \ll 1 \), etc. We then get an estimate \( \delta c^2 = O(e_2 \mu^2) = O(\sqrt{\varepsilon \mu^2}) \ll m^2 \). As in the superrenormalizable case, the speed term can be naturally much smaller than the gap scale. This prediction can be verified by a direct loop calculation. The leading contribution to \( \delta c^2 \) comes from several two-loop diagrams, including \( \frac{1}{\Lambda^2} \) with one \( e_2 \) vertex. Each loop in this diagram is separately linearly divergent, giving \( \delta c^2 \sim e_2 \Lambda^2 = O(\sqrt{\varepsilon \mu^2}) \), in accord with our scaling argument.

The story extends naturally to the broken phase of the renormalizable LSM, although this theory is technically rather complicated: The \( \langle \phi \rangle \) itself is no longer given by \([4]\) but it is at the minimum of a generic fifth-order polynomial in \( \phi^I \phi^I \). It is thus more practical to run our argument directly in the low-energy NLSM. The advantage is that even for the generic renormalizable LSM \([3]\), the leading-order NLSM action is of the general form \([2]\). The leading order of matching gives \( G = 1/v \), with \( v \) the radius of the vacuum manifold \( S^{N-1} \). The NLSM is weakly coupled when this radius is large. The enhanced “quadratic shift” symmetry of the NG modes \( \pi^4 \) in their free-field limit implies \( G^2 = O(\varepsilon/\mu) \) and \( c^2 = O(\varepsilon \mu^2) \) with \( \lambda_{1,2} = O(1) \), and predicts

\[
e^2 = O(G^2 \mu^3).
\]

The naturalness scale \( \mu \) is set by the gap of the \( \sigma \) particle, which is generally of order \( m \). Thus, \([12]\) implies that in the large-\( v \) regime of the weakly-coupled NLSM, the speed term is naturally much smaller than the naturalness scale. This can be again confirmed by a direct loop calculation: The leading contribution to \( \delta c^2 \) comes from the one-loop diagram \( \overline{\dots} \). This diagram is cubically divergent and its vertex gives a \( G^2 \) factor, leading to \( \delta c^2 \sim G^2 \Lambda^3 \). Setting \( \Lambda \sim \mu \) confirms our scaling prediction \([12]\).

In the special case of our super-renormalizable LSM, we can go one step further, and use \([4]\) and \( G = 1/v \) to reproduce again our earlier result, \( \delta c^2 = O(\lambda / m) \).

**Discussion:** We have shown that Type-A NG modes can naturally have an anomalously slow speed, characterized by an effective \( z = 2 \) dispersion relation. Our arguments can be easily iterated, leading to Type-A NG modes with higher dispersion of \( z = 3, 4, \ldots \). In such higher multicritical cases, the smallness of all the relevant terms is protected by the enhanced “polynomial shift” symmetry in the free-field limit, with \( a^I \) now a polynomial in \( x^I \) of degree \( 2z - 2 \).

At fixed spatial dimension, this pattern of multicritical symmetry breaking will eventually run into infrared divergences and a multicritical version of the Coleman-Mermin-Wagner theorem \([33, 36]\). We can increase \( z \) until we reach \( z = D \), at which point no symmetry breaking with this or higher scaling is possible – the candidate NG mode described by the free \( z = D \) scalar in \( D + 1 \) spacetime dimensions does not exist as a physical object,
since its propagator is a log and depends on the infrared regulator. This theory would have to supply its own infrared regulator, for example, by crossing over to $z < D$ in the far infrared, after spending a lot of RG time in the vicinity of $z = D$ at intermediate scales.

Our results also extend easily to Type-B NG modes, which break time reversal invariance. Instead of their generic $z = 2$ dispersion, they can exhibit a $z = 4$ (or higher) behavior over a large range of energy scales.

In all these cases, the multicoercivity of the NG modes will have consequences for their low-energy scattering, generalizing the low-energy theorems known from the relativistic case [7]. The scattering amplitudes will exhibit a higher-power effective dependence on the momenta, with the power controlled by $z$.

Finally, it would be very interesting to extend our analysis to the spatially modulated phases of Lifshitz theories, in which the spacetime symmetries are further broken spontaneously, and where one can expect spatially modulated NG modes.

The results of this paper refine the classification of NG modes in non-relativistic systems, and we expect them to be useful for understanding symmetry breaking in a broad class of phenomena, including relativistic matter at nonzero density or chemical potential, and areas of condensed matter, such as superconductivity, quantum critical phenomena and dynamical critical systems. Since our results also shed interesting new light on the concept of Naturalness, we are hopeful that they may stimulate new insights in areas where puzzles of Naturalness have been most prominent: particle physics, quantum gravity and cosmology.

Acknowledgements: We thank Chien-I Chiang, Charles Melby-Thompson and Zachary Stone for useful discussions. This work has been supported by NSF Grant PHY-1214644, DOE Grant DE-AC02-05CH11231, and by Berkeley Center for Theoretical Physics.

[33] J. Zinn-Justin, Quantum field theory and critical phenomena, Ch. 30.6 (Oxford University Press, 2002).