Cosmic strings and the topological theory of defects

Author: Leonardo García Heveling
*Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.**

Advisor: Oleg Bulashenko

Abstract: A mathematical characterization of topological defects is provided, with a discussion of topological stability and the necessary conditions for stable defects to appear. Defect formation in symmetry-breaking phase transitions of the Universe is discussed, and a simple model for the formation of cosmic strings is given. Geodesics in a cosmic string spacetime are calculated for the study of gravitational lensing phenomena.

I. INTRODUCTION

When a physical system undergoes a symmetry-breaking phase transition, it might be left with topological defects. Their study started in the mid-1970's, in both condensed matter physics [1] and cosmological phase transitions [2]. As we shall see, topology, and in particular, homotopy theory, provides a very natural formulation of the theory of defects. It enables us to classify configurations of a physical system in terms of topological invariants, with interesting physical implications. Different physical systems can be related if their topology is the same. For example, cosmic strings in the Universe are related to disclinations in solids. This way, different areas of physics can be described by the same mathematical structures, and experiments in one area can be related to the other. A mathematical description of topological defects is given in sec. II. Applications to phase transitions in the Universe, and the relationship between topological defects and symmetry-breaking, in sec. III. In sec. IV, a taste of interesting physics due to defects is given by the study of cosmic strings, with emphasis on gravitational lensing phenomena that could lead to observations.

II. TOPOLOGICAL DEFECTS

A. Ordered media

Let us first give some definitions based on [1]. An *ordered medium* is a region of physical space with a function $f(\mathbf{r})$, the *order-parameter field*, that assigns to every point of the medium an *order parameter*. The order parameter is an element of some topological space, the *order-parameter space*, that depends on the physical problem. In general, the lowest energy configuration will be such that $f(\mathbf{r})$ is continuous. However, we will be most interested in those cases where $f(\mathbf{r})$ is continuous everywhere except in a lower dimensional subregion of the ordered medium, like a point, line or surface. We will call this subregion a *topological defect*.

*Electronic address: `leogh_33@hotmail.com`

If it is possible to eliminate the defect by making purely local changes to the order-parameter field, the defect is considered to be *topologically unstable*. In other words, we only need to change the definition of $f(\mathbf{r})$ in an arbitrarily small neighborhood of the unstable defect to eliminate it. On the contrary, some defects can only be eliminated by altering the order-parameter field in a large region, even the entire ordered medium. These defects are called *topologically stable*. However, to study the physical stability of a defect, it is also necessary to make the relevant energy considerations [1].

B. An example: planar spins

In the following example, the ordered medium is the two-dimensional plane and the order-parameter space is the unit circle S^1. The order-parameter field $f(x, y)$ now assigns to every point on the plane a two-dimensional unit vector of S^1, as in fig. 1. Let us consider a point P and a circular contour $C(\lambda)$ on the plane, centered at P, parametrized by λ. We assume that $f(x, y)$ is continuous everywhere, except maybe at P. If we move along the entire contour C, once we get back to the starting point, $f(C(\lambda))$ will have turned a certain angle $2\pi n$ along the unit circle S^1. n is the *winding number*, and has to be an integer if $f(x, y)$ is continuous along C, as we assumed. We consider n positive if $f(C(\lambda))$ turns around the unit circle in the same manner (clockwise or counterclockwise) as $C(\lambda)$ around P, and negative in the opposite case.

If we now take the circular contour C and enlarge (or shrink) its radius, the winding number n must remain the same, because being an integer, it can only change discontinuously. This means that if $n \neq 0$, $f(x, y)$ must be singular at P, because for any arbitrarily small radius, n is the same. In other words, there is a topological defect at P. Further, for any arbitrarily large radius, n is also the same. Since we just pointed out that only a configuration with $n = 0$ can be continuous at P, the defect is stable, because we would need to make changes arbitrarily far from P to eliminate it.

n is an example of *topological charge*, a discrete quantity associated to a particular order-parameter field configuration. Topological charges are invariant under continuous local transformations of the order-parameter

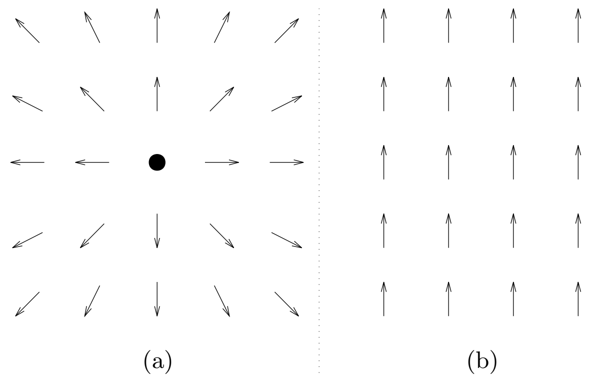

FIG. 1: (a) Configuration of planar spins with a defect of winding number $n = 1$ (black dot). (b) Configuration with $n = 0$ and no defects.

field. In this sense, it can also be proven that any singularity with $n = 0$ is topologically unstable, and that one defect can be replaced by another one of the same n making only local changes [1]. Also, in a configuration with two defects, the winding number of a contour encircling them both is the sum of the winding numbers for each defect. This implies that two defects of opposite winding number can annihilate each other, meaning that if they are close together, we can eliminate them making changes only in a neighborhood that contains them both. The planar spins example can be expanded to a 3-dimensional ordered medium, and then the defects are not points, but lines.

C. Homotopy theory

We will now introduce some more mathematics, favoring intuition over rigor. A more formal treatment is found in [3]. Our ordered medium is now $\mathbb{R}^3$, with some arbitrary order-parameter space V and an order-parameter field $f : \mathbb{R}^3 \to V$. Let S^n be the n-dimensional sphere, with $n \leq 2$. This way S^2 is the ordinary sphere, S^1 a circle and S^0 a set of two points with $|\mathbf{x}| = 1$. We will take freedom in choosing the origin of coordinates. The restriction of f to S^n provides a mapping of the n-sphere to order-parameter space $f|_{S^n}(\mathbf{x}) : S^n \to V$. If we consider two different configurations f,g (order-parameter fields to the same V), then we have two such mappings $\lambda = f|_{S^n}$, $\sigma = g|_{S^n}$. λ and σ are said to be *homotopic* if there exists another mapping $H(t, \mathbf{x}) : [0, 1] \times S^n \to V$, that is continuous in $\mathbf{x} \in S^n$ and $t \in [0, 1]$, and such that $H(0, \mathbf{x}) = \lambda(\mathbf{x})$ and $H(1, \mathbf{x}) = \sigma(\mathbf{x})$. H is then called an *homotopy*. t has to be regarded as a parameter (not to be confused with the order-parameter nor the parametrization of S^n). The property of two mappings being homotopic defines an equivalence relation.

We denote by $\pi_n(V)$ the set of all equivalence classes of homotopic maps from S^n to V, also called homotopy classes. An element of $\pi_n(V)$ is therefore a family (homotopy class) of mappings that are all homotopic among each other, and not homotopic to any mapping from a different class. $\pi_n(V)$ has a group structure for $n \geq 1$ and is called the *nth*-homotopy group, but this is not important for our purposes. The particular structure of $\pi_n(V)$, for each n, depends on the topology of V. If $\pi_n(V)$, for some n, contains only one homotopy class, then it is said that "the *nth*-homotopy group of V is trivial" and denoted by $\pi_n(V) = 0$. When this is the case, all mappings from S^n to V are homotopic to each other.

Now we can provide a theorem that states a necessary condition for the existence of stable topological defects of different dimensionality. We will call them point-, line- and wall defects if their dimensionality is 1, 2, and 3 respectively. The theorem states that there can be stable:

1. Point defects only if $\pi_2(V) \neq 0$.
2. Line defects only if $\pi_1(V) \neq 0$.
3. Wall defects only if $\pi_0(V) \neq 0$.

A rigorous proof is provided by [3]. We will see a brief sketch of the proof for the case of point defects. Let us consider a point defect at P and a 2-sphere S^2_ρ of arbitrarily small radius ρ, centered at P. If $\pi_2(V) = 0$, whatever mapping $f|_{S^2_\rho}$ provides, it will be homotopic to a constant map. Therefore a homotopy $H(t, \mathbf{x})$ exists, with $H(1, \mathbf{x}) = f|_{S^2_\rho}(\mathbf{x})$ and $H(0, \mathbf{x}) = C$. This way, we can redefine the order-parameter field inside S^2_ρ as $f'(r, \theta, \phi) = H(\frac{r}{\rho}, (\theta, \phi))$. Here we have used spherical coordinates, and associate the homotopy-parameter with the radial coordinate to ensure continuity as we approach the origin. This new order-parameter field is continuous everywhere, and we only had to make changes inside a region of arbitrarily small radius ρ, so the point defect at P had to be unstable.

This classification of defects in $\mathbb{R}^3$ can be expanded to include a fourth type, called *texture* [3] [4], associated with non-trivial $\pi_3(V)$. However, textures are not defects according to our definition, because the order-parameter field need not be singular anywhere. Instead, they appear when the entire order-parameter field (in all of $\mathbb{R}^3$), is not homotopic to a constant order-parameter field (which is regarded as the lowest energy configuration).

III. DEFECT FORMATION IN THE EARLY UNIVERSE

A. Spontaneous symmetry breaking

It is a typical feature of modern cosmological theories to describe the evolution of the universe through various phase transitions, that occur as the universe cools down after the Big Bang [4]. These phase transitions are governed by some set of scalar fields $\boldsymbol{\varphi} = \{\phi^a\}$. They obey a Lagrangian density of the form [5]

$$\mathscr{L} = K(\partial^\mu \boldsymbol{\varphi}) - V(\boldsymbol{\varphi}) \tag{1}$$

with a kinetic term K that depends on the field derivatives and a (classical or effective) potential V that depends on the fields. The phase transition has an associated critical temperature T_c. For $T > T_c$, V has an absolute minimum at $\boldsymbol{\varphi} = 0$, so in the vacuum state, the expectation value is $\langle\boldsymbol{\varphi}\rangle = 0$. However, for $T < T_c$, there is a manifold of degenerate minima, the vacuum manifold $\mathcal{M}$. The field will then have a vacuum expectation value (vev) $\langle\boldsymbol{\varphi}\rangle = \boldsymbol{\varphi}_0 \in \mathcal{M}$ that is chosen randomly, so there is a spontaneous symmetry breaking.

The correlation length for the vev cannot be greater than the causal horizon. Therefore, causally disconnected regions (domains) of space generally have different values of the vev. When these domains come into causal contact, the kinetic term in (1) will make the field "smooth out" to lower its derivatives. Ideally, the field will end up settling to a uniform value, so that $K(\partial^\mu \boldsymbol{\varphi}) = 0$. However, if some homotopy group of $\mathcal{M}$ is non-trivial, it can happen that the field reaches a configuration with stable topological defects (II C). These defects cannot be eliminated by thermal fluctuations, so they will remain. This is the Kibble mechanism for defect formation [2]. Here, $\mathcal{M}$ plays the role of order-parameter space. A simple example is provided in sec. IV A.

Using Lie groups allows us to get a deeper understanding of spontaneous symmetry breaking and its relation to the vacuum manifold $\mathcal{M}$. Let G be the symmetry group of our model. An element $g \in G$ acts through a representation D of G as $\boldsymbol{\varphi} \to D(g)\boldsymbol{\varphi}$, in a way that leaves (1) invariant. When $T > T_c$, the expectation value $\langle\boldsymbol{\varphi}\rangle = 0$ is also invariant, because it is zero. However, $\boldsymbol{\varphi}_0$ is only invariant under the action of a subgroup $H \subset G$. If we take an arbitrary vev $\boldsymbol{\varphi}_0$, all other possible vevs can be constructed as $D(gh)\boldsymbol{\varphi}_0$, for some $g \in G$ and $\forall h \in H$. This is because $D(gh) = D(g)D(h)$ and $D(h)\boldsymbol{\varphi}_0 = \boldsymbol{\varphi}_0$, and allows us to establish

$$\mathcal{M} = G/H \tag{2}$$

where G/H is the coset space. This is important for two reasons. First, it allows us to relate $\mathcal{M}$ to the symmetry breaking we are dealing with. Second, calculating homotopy groups of coset spaces can be made easier by the use of universal covers and exact sequences [5].

B. Physical implications

Taking another look at the Lagrangian density (1) reveals that topological defects will have an associated energy. Near a defect, the field derivatives ideally grow without bound. Consider, for example, the planar spins of II B, where the unit vector has to turn "faster" around the unit circle as we consider circular contours closer to the defect. This gives the defect a "kinetic" energy through $K(\partial^\mu \boldsymbol{\varphi})$. Also, we do not expect the field to be truly singular at the defect, as this would be unphysical. Instead, close to the defect, the field takes values that are not on the vacuum-manifold, thus having a higher potential energy. This energy justifies regarding defects as real physical entities, rather than just mathematical features of a particular model.

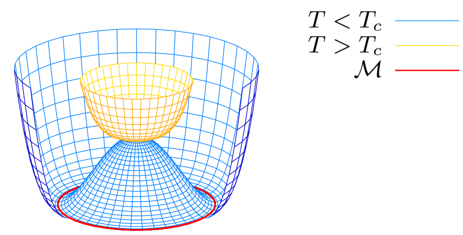

FIG. 2: In blue, the Mexican hat potential, with its vacuum manifold in red. In orange, the potential for the phase of unbroken symmetry.

In the cosmological context, point defects are called *monopoles* and line defects *cosmic strings*. Monopoles are expected to have a net magnetic charge. Their existence would justify the quantization of electric (and magnetic) charge, as pointed out by Dirac. Cosmic strings are a candidate for the explanation of structure formation in the universe and CMB fluctuations [4]. Wall defects and textures are also possible in some models, but their physical relevance is somewhat more controversial [6]. The impact of defects on the universe, and the possibility of observing them today, depends on their dynamics and interactions (see [5] for a detailed study). As of yet, no observations have been made.

IV. COSMIC STRINGS

A. A simple model of string formation

We will now consider a very crude, classical model with one complex scalar field ϕ, that is nevertheless enough to illustrate cosmic string formation [5]. In the broken symmetry phase ($T < T_c$), the Lagrangian density is

$$\mathscr{L} = (\partial_\mu \phi^*)(\partial^\mu \phi) - \frac{\lambda}{4}(\phi^*\phi - \eta^2)^2 \tag{3}$$

where λ, η are positive constants, and the potential has a Mexican hat shape (fig. 2). It is invariant under the $U(1)$ group of global phase transformations $\phi(x) \to e^{i\alpha}\phi(x)$ (global in the sense that α is the same $\forall x$). However, the minima of the potential lie on the circle $\phi^*\phi = \eta^2$. Therefore we have a vev $\langle\phi\rangle = \eta e^{i\theta}$, so the ground state is not invariant under $U(1)$ and the symmetry is spontaneously broken. We assume here that for $T > T_c$, the potential has a shape that is also invariant under $U(1)$, and a single absolute minimum at $\phi = 0$.

Since there are no elements other than the identity in $U(1)$ that leave $\eta e^{i\theta}$ invariant, the symmetry is completely broken and $\mathcal{M} = U(1)$. $U(1)$ has the topology of the S^1 circle, and $\pi(S^1) = \mathbb{Z}$ [3], so line defects can appear. In fact, the first homotopy group is the integers because each homotopy class can be characterized by its winding number, as in sec. II B.

B. The straight string metric

The simplest type of cosmic string is a straight string that extends along the z-axis. It can be described by the energy-momentum tensor

$$T_\mu^\nu = \sigma(r)\mathrm{diag}(1,0,0,1) \tag{4}$$

The simplest model is to consider $\sigma(r) = \mu\delta(x)\delta(y)$, where μ is the mass per unit length, as in [5]. Gott solves a slightly more sophisticated case in [7], where the energy-momentum tensor is non-zero inside a small radius around the string. Both in this case and for the delta distribution, the (outside) metric is

$$ds^2 = dt^2 - dz^2 - dr^2 - \beta^2 r^2 d\varphi^2 \tag{5}$$

where β can be related to the mass per unit length μ and gravitational constant G, as $\beta = 1 - 4G\mu$. Redefining the angular coordinate as $\phi = \beta\varphi$, the metric takes the Minkowski form locally

$$ds^2 = dt^2 - dz^2 - dr^2 - r^2 d\phi^2 \tag{6}$$

but not globally, because now $-\pi + \Delta < \phi \leq \pi - \Delta$, where we have defined the angle deficit

$$\Delta = (1-\beta)\pi = 4\pi G\mu \tag{7}$$

C. Geodesics

In this subsection we calculate the trajectories (the spatial images of geodesics) of light and particles around a straight cosmic string. We will ignore the z-coordinate, considering only the trajectories that lie in a plane perpendicular to the string, which has the geometry of a cone. This is equivalent to considering a point-mass in 2+1-dimensional gravity [8].

Using the coordinates and metric (5), the geodesics are best calculated using the Hamiltonian $2\mathscr{H} = g_{\mu\nu}\dot{q}^\mu\dot{q}^\nu$, where q^μ are the coordinates and the dot represents differentiation with respect to the affine parameter. Hamilton's equations give us two conserved quantities

$$\begin{aligned} \dot{t} &= E \\ -\beta^2 r^2 \dot{\varphi} &= L \end{aligned} \tag{8}$$

For light-like geodesics, we have $2\mathscr{H} = 0$. Using this, we find

$$\dot{r}^2 = E^2 - \frac{L^2}{\beta^2 r^2} \tag{9}$$

This equation is equivalent to a one dimensional Newtonian mechanics problem with an effective potential $\frac{L^2}{\beta^2 r^2}$.

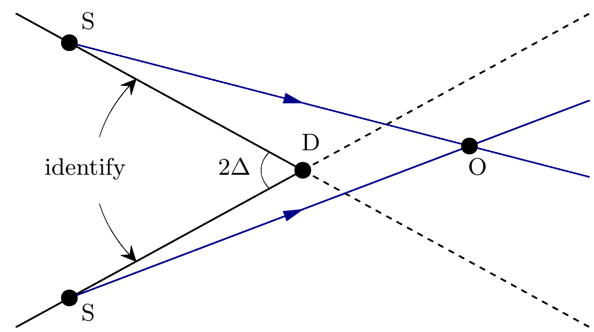

FIG. 3: Two distinct geodesics from a single source S to an observer O behind the cosmic string D. The missing wedge is the area between the two solid lines, which are identified. The area between the dashed lines containing the observer is the double imaging sector. This representation corresponds to the (r, ϕ) coordinates.

There is one turning point

$$r_0 = \frac{L}{\beta E} \tag{10}$$

We can now take the square root of (9) and divide by $\dot{\varphi}$ to find the orbital equation

$$\frac{dr}{d\varphi} = \pm\beta r^2\sqrt{\frac{1}{r_0^2} - \frac{1}{r^2}} \tag{11}$$

where the square root can take both the positive and the negative determination. Solving it gives us an equation for the trajectories

$$r\cos\beta(\varphi - \varphi_0) = r_0 \tag{12}$$

where φ_0 is an integration constant and can be interpreted as the angle of maximum approach of the geodesic to the string. This result was already obtained in [8].

For time-like geodesics, $2\mathscr{H} = 1$, and the whole process is analogous, giving us the same equation (12), but now with

$$r_0 = \frac{L}{\beta\sqrt{E^2 - 1}} \tag{13}$$

The fact that light and massive particles follow the same trajectories, albeit at different speeds, is even clearer when using the alternative angular coordinate $\phi = \beta\varphi$, that makes the metric to be Minkowski-like (6). This way the trajectories are straight lines that live on a plane with a missing wedge of angle 2Δ (fig. 3). The two boundaries of the missing wedge are identified, i.e. they represent the same physical points in space.

D. Gravitational lensing

Let us now consider a situation with a light source at (r_1, π) and an observer at (r_2, φ_2). We have to im-

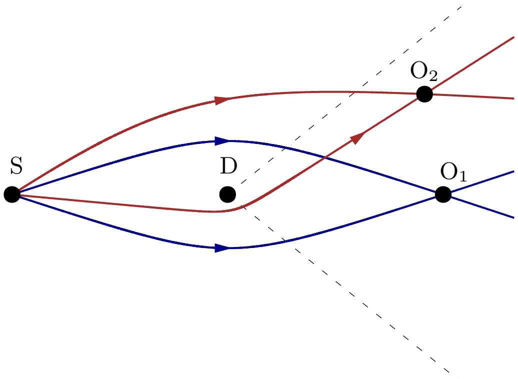

FIG. 4: A situation with source S, a cosmic string D and two observers O_1, O_2. Each observer sees two distinct images of S. The area between the dashed lines containing the observers is the double imaging sector. This representation corresponds to the (r, φ) coordinates.

pose that (12) is valid at both points. This was done numerically with Maple to generate fig. 4. For a qualitative analysis we use the change of coordinates $\phi = \beta\varphi$. We see in fig. 3 that there are two distinct trajectories that connect the source and the observer if and only if $-\Delta < \phi < \Delta$. This region is the *double imaging sector*; an observer here will see two distinct images of the source. This is remarkable because observations of double imaging could indicate the presence of a cosmic string.

In the (r, φ) coordinates, the trajectories are given by (12) and shown in fig. 4. The double imaging sector is now $-(\frac{1}{\beta} - 1)\pi < \varphi < (\frac{1}{\beta} - 1)\pi$, having in mind that Δ and β are related by (7).

Another interesting situation is to consider a photon coming in from infinity with a certain impact parameter a. In this case, the initial conditions can be imposed analytically in the (r, φ) coordinates. If the particle comes from the π direction, its trajectory will tend asymptotically to a straight line with equation $a = r\cos(\varphi - \frac{\pi}{2})$. This allows us to establish

$$\begin{aligned} r_0 &= \beta a \\ \varphi_0 &= \pi(1 - \frac{1}{2\beta}) \end{aligned} \tag{14}$$

This result is of course also valid for massive particles, and it is worth mentioning that it is independent on the energy or velocity of the particle. Cosmic strings do not cause gravitational attraction [5], and again changing to the (r, ϕ) coordinates with metric (6) makes this clear. Another result derived of (14) is that the deflection angle does not depend on the impact parameter, and it is simply $\varphi_D = (1 - \frac{1}{\beta})\pi$. This last result is important for a more throughout study of gravitational lensing, which is necessary to make observational predictions.

The results of this section can also be applied to the study of light propagation in materials with disclinations [9]. These are line defects that can be described by the same metric as cosmic strings, also extending to the case $\beta > 1$. This allows to conduct experiments in the lab that can be related to what happens in the Universe.

V. CONCLUSIONS

We have seen that the possibility of topological defects being created in a phase transition only depends on the symmetry being broken. If the right homotopy group is non-trivial, defect formation by the Kibble mechanism is practically unavoidable. The study of what impact these defects can have on the evolution of the Universe, and whether they might still exist today, is beyond our scope and depends on the particular model (see [5] for more). Be that as it may, observations of defects would be of huge importance not only to cosmology, but also particle physics, as they would confirm theories that allow them and disregard those that do not. Gravitational lensing provides one possible means of observation of cosmic strings.

Acknowledgments

I want to thank my advisor Oleg Bulashenko for interesting me in a subject I hardly knew existed, and Carles Casacuberta for teaching me topology. I also want to thank my friends and family, especially my father who recommended me ref. [3].

[1] Mermin, N.D., "The topological theory of defects in ordered media". Rev. Mod. Phys. **51**(3): 591-648 (1979).

[2] Kibble, T.W.B., "Topology of cosmic domains and strings", J. Phys. A.: Math. Gen. **9**(8): 1387-1398 (1976).

[3] Nash, C. and Sen, S., *Topology and Geometry for Physicists*, (Academic Press, 3rd ed., 1987).

[4] Peacock, J.A., *Cosmological Physics*, (Cambridge University Press, 9th ed., 2010).

[5] Vilenkin A. and Shellard E.P.S., *Cosmic strings and other topological defects*, (Cambridge University Press, 1st ed., 1994).

[6] Gorbunov, D.S., Rubakov, V.A., *Introduction to the Theory of the Early Universe*, (World Scientific Publishing, 1st ed., 2011).

[7] Gott, J.R., III, "Gravitational lensing effects of vacuum strings: exact solutions". ApJ **288**: 422-427 (1985).

[8] Deser, S. and Jackiw, R., "Classical and Quantum Scattering on a Cone". Commun. Math. Phys. **118**: 495-509 (1988).

[9] Yong-Lian Zhang et. al., "Steering electromagnetic beams with conical curvature singularities". Optics Letters **40**(20), 4783-4786 (2015).