Is string field theory background independent?

Matěj Krátký[*], Bhanu Narra[†], & James Read[‡]

Abstract

String field theory is supposed to stand to perturbative string theory as quantum field theory stands to single-particle quantum theory; as such, it purports to offer a substantially more general and powerful perspective on string theory than the perturbative approach. In addition, string field theory has been claimed for several decades to liberate string theory from any fixed, background spatiotemporal commitments—thereby (if true) rendering it 'background independent'. But is this really so? In this article, we undertake a detailed interrogation of this claim, finding that the verdict is sensitive both to one's understanding of the notion of background independence, and also to how one understands string field theory itself. Although in the end our verdicts on the question of the background independence are therefore somewhat mixed, we hope that our study will elevate the levels of systematicity and rigour in these discussions, as well as equip philosophers of physics with a helpful introduction to string field theory and the variety of interesting conceptual questions which it raises.

Contents

[*]matej.kratky@etu.unige.ch
[†]bhanu.narra@pmb.ox.ac.uk
[‡]james.read@philosophy.ox.ac.uk

1 Introduction

Background independence is a (supposed) theoretical virtue that, as the name suggests, aims to capture whether a given physical theory depends on or involves the choice of any fixed background structure. However, despite there being clear paradigmatic cases of both background independent and background dependent theories, stating the precise definition of background independence has proved rather challenging. One (if not *the*) paradigmatic case of a background independent theory is general relativity, which can be formulated on arbitrary Lorentzian manifolds dynamically coupled to stress-energy content via the Einstein equations and which hence, intuitively, does not involve a choice of fixed background structure. On the other hand, special relativity, at least in its geometrical formulation, requires Minkowski spacetime—clearly a fixed background structure—and hence is considered background dependent.[1] Clarifying the definition of background independence is particularly urgent in light of the authority it has acquired in quantum gravity circles.[2]

In this article, we consider the background independence of string field theory (SFT), one such candidate for a theory of quantum gravity. In the first approximation, one can think of the relationship between SFT and perturbative string theory as being analogous to the relationship between relativistic QFT and single-particle quantum mechanics.[3] While perturbative string theory describes vibrational excitations of a single relativistic string, SFT aims to describe excitations of a *string field*. In assessing the background independence

[1] There are subtleties here regarding different ways of formulating special relativity, but we'll elide them. For further discussion, see Pooley (2017) and Read (2023).

[2] Fortunately, there is by now a lot of philosophical work that aims to fill this gap (see, e.g., Read (2023)).

[3] Sometimes physicists call SFT the *second-quantized* version of perturbative string theory.

of SFT, we draw heavily on the physics literature. In particular, we discuss (a) the series of mathematical results by Sen and Zwiebach (1994a,b) and Sen (2018) that allegedly demonstrate the background independence of SFT, (b) the 'manifestly background independent' formulation of SFT due to Witten (1992), and (c) a contemporary extension by Ahmadain et al. (2024). We highlight some limitations of these mathematical results, and evaluate the background independence of SFT with respect to the various definitions of background independence given by Read (2023). In the end—as one might expect!—we find that the question of whether SFT is background independent depends both on how one understands 'background independence' and on how one formulates SFT.

Here's our plan for the article. In §2, we provide the necessary technical background on worldsheet (perturbative) string theory; in §3, we do the same for (super)string field theory. In §4, we review arguments for SFT's background independence, and in §5 we present the structures which are invariant between SFTs. Doing the latter will be crucial for a full assessment of the background independence of SFT—a task we take up in §6.

2 Worldsheet string theory

In this section, we introduce worldsheet string theory as a stepping stone to our subsequent discussion of SFT. We follow the conformal field theory (CFT) approach to worldsheet string theory (see Polchinski (1998)) that is best suited to understanding the formulation of covariant SFT. In §2.1, we provide the relevant background on the physics of classical strings situated in some ambient 'target space' with background fields. In §2.2, we review relevant details of thinking about string theory as a 2D CFT on the worldsheet.

2.1 Classical relativistic strings in background fields

The starting point for worldsheet string theory is the study of classical relativistic string propagating in Minkowski spacetime. Let $(\mathbb{R}^D, \eta)$ be a D-dimensional Minkowski spacetime which we call *target space* and let Σ_g be a Riemann surface of genus g which we call the *worldsheet*. Suppose there is an embedding $X : \Sigma_g \to \mathbb{R}^D$ which may also be specified by D independent functions $X^\mu : \Sigma_g \to \mathbb{R}$ where μ runs from 0 to D and labels global coordinates in target space. Any such embedding describes a kinematically possible evolution of the vibrating classical string in target space. The dynamically possible embeddings are then picked out by equations of motion which follow from the Polyakov action,

$$S_{\mathrm{P}}[X^\mu, h_{ab}] = -\frac{1}{4\pi}\int_{\Sigma_g} d^2\sigma\ \sqrt{-h}h^{ab}\partial_a X^\mu \partial_b X^\nu \eta_{\mu\nu}, \tag{1}$$

where h_{ab} is a Lorentzian metric on the worldsheet metric and $h := \det(h_{ab})$ (h^{ab} is its inverse). The Polyakov action is highly symmetric: it is invariant under both worldsheet reparametrizations and Weyl rescalings of the worldsheet

metric. More explicitly, if $f : \Sigma_g \to \Sigma_g$ is a diffeomorphism and if $\omega \in C^\infty(\Sigma_g)$ is a smooth function, then

$$S_{\mathrm{P}}[f^* X^\mu, f^* h_{ab}] = S_{\mathrm{P}}[X^\mu, h_{ab}] \quad \text{and} \quad S_{\mathrm{P}}[X^\mu, e^{2\omega} h_{ab}] = S_{\mathrm{P}}[X^\mu, h_{ab}] \tag{2}$$

express invariance under worldsheet reparametrization and Weyl rescalings, respectively.

Note that S_{P} already involves a choice of background $(\mathbb{R}^D, \eta)$. However, different background choices are possible. For instance, one can ask how classical relativistic strings propagate on general Lorentzian manifolds (M, g). Once again, suppose that we have an embedding $X : \Sigma_g \to M$. The natural generalization of S_P is

$$S_1[X^\mu, h_{ab}] = -\frac{1}{4\pi} \int_{\Sigma_g} d^2\sigma \ \sqrt{-h} h^{ab} \partial_a X^\mu \partial_b X^\nu g_{\mu\nu}(X), \tag{3}$$

where $g_{\mu\nu}$ are the metric components.[4]

Following the lead of classic texts such as Green et al. (2012), we also consider further generalizations of S_{P} which couple the string to the remaining massless fields of the closed string spectrum: the two-form gauge field known as the Kalb–Ramond field and the scalar dilaton field. The action terms corresponding to these couplings are respectively

$$S_2[X^\mu, h_{ab}] = -\frac{1}{4\pi} \int_{\Sigma_g} d^2\sigma \ \epsilon^{ab} \partial_a X^\mu \partial_b X^\nu B_{\mu\nu}(X), \tag{4}$$

$$S_3[X^\mu, h_{ab}] = \frac{1}{4\pi} \int_{\Sigma_g} d^2\sigma \ \sqrt{-h} \Phi(X) R, \tag{5}$$

where ϵ^{ab} is the antisymmetric tensor density, R is the worldsheet Ricci scalar, and Φ and $B_{\mu\nu}$ are some fixed background choices for the dilaton and Kalb–Ramond field respectively. Since the combination $d^2\sigma\sqrt{-h}$ is diffeomorphism invariant, the diffeomorphism invariance of S_1, S_2, and S_3 is manifest. Weyl invariance is manifest for S_1 and S_2 but S_3 in general will not be Weyl invariant. More precisely, S_3 is classically invariant only under global Weyl transformations. However, quantum mechanically we can force it to be locally Weyl invariant by forcing the beta function of the theory to vanish, making the theory a CFT, as discussed below. Local Weyl invariance is critical for the theory to

[4] There also exists a coordinate-free definition of the Polyakov action; however, its explanation involves some technicalities. Let $[\Sigma_g, M]$ be the space of smooth maps from Σ_g to M and let $\mathrm{Met}(\Sigma_g)$ be the space of metrics on Σ_g. Associated to the embedding $X : \Sigma_g \to M$ is the tangent map $T_p X : T_p\Sigma_g \to T_{X(p)}M$ at every $p \in \Sigma_g$ and also the map of tangent bundles $TX : T\Sigma_g \to TM$. Because both Σ_g and M come equipped with Lorentzian metrics, we may define the *adjoint* of $T_p X$ as the map $T_p X^* : T_{X(p)}M \to T_p\Sigma_g$ such that for any two vectors $v \in \Gamma(T_p\Sigma_g)$ and $w \in \Gamma(T_{X(p)}M)$ the relation $h_p(v, T_{X(p)}X(w)) = g_{X(p)}(T_p X(v), w)$ holds. One may then define the norm $|T_p X|^2 = \mathrm{Tr}(T_p X^* \circ T_p X)$ at p and the *global function* on Σ_g given by $|TX|^2(p) = |T_p X|^2$. Moreover, the metric h also induces a unique volume-form $\mathrm{vol}(h)$ on Σ_g so that the Polyakov action may be defined as a functional $S_P : [\Sigma_g, M] \times \mathrm{Met}(\Sigma_g) \to \mathbb{R}$ given by $S_P[X, h] = \int_{\Sigma_g} |TX|^2 \mathrm{vol}(h)$. For more details, see Dolgachev (1999, Lecture 4).

be consistent, and so we find that classical string backgrounds correspond to solutions of the beta function equations, which in particular contain the Einstein field equations.[5] We'll discuss the more general situation for perturbative string theory in §2.2.3, and we'll discuss the situation for SFT in §3.5.

2.2 Some details about the worldsheet

We now review some details about worldsheet string theory—i.e., string theory understood as a 2D CFT—which will be relevant for the construction of covariant string field theories in the next section.

2.2.1 The Hilbert space

Our starting point will again be the Polyakov action in flat spacetime, (1). When viewed as a two dimensional theory on Σ_g, the symmetries (2) are treated as gauge symmetries, and a gauge fixing procedure must be implemented.[6] An important consequence of the gauge fixing is the introduction of anticommuting Faddeev–Popov *ghost fields* b^{ab} and c_a, with action

$$S_{\text{gh}} = \frac{1}{2\pi}\int_{\Sigma_g} d^2\sigma\sqrt{-h}b^{ab}\nabla_a c_b. \tag{6}$$

In the quantization, one finds that the symmetries (2) will be anomalous unless $D = 26$. Since these are local symmetries, the theory is consistent only if $D = 26$. Furthermore, the gauge fixing leaves behind a residual symmetry group called the *conformal group*, which is an extension of the Poincaré group that includes dilatations (scale transformations). This makes the theory a *conformal field theory* or CFT.

The gauge fixing reduces the naïve Hilbert space of the theory (which is a product of the Polyakov and ghost Hilbert spaces) down to the gauge invariant states:

$$\mathcal{H}_0 = \mathcal{H}_{\text{P}} \otimes \mathcal{H}_{\text{gh}} \to \mathcal{H}_{\text{phys}}. \tag{7}$$

This reduction is implemented by introducing an operator Q_B, called the *BRST operator*, which acts on $\mathcal{H}_0$. The BRST operator is nilpotent ($Q_B^2 = 0$), and the cohomology classes of this operator give the states in $\mathcal{H}_{\text{phys}}$. These are the states in $\mathcal{H}_0$ that satisfy

$$Q_B\,|\psi\rangle = 0, \tag{8}$$

with the equivalence

$$|\psi\rangle \sim |\psi\rangle + Q_B\,|\lambda\rangle\,,\ |\lambda\rangle \in \mathcal{H}_0. \tag{9}$$

It will be easier to work with the larger Hilbert space $\mathcal{H}_0$ while keeping in mind the equivalence (9).

[5] This reasoning has been discussed by philosophers in Huggett and Vistarini (2015), Huggett and Wüthrich (2025), and Read (2019).

[6] See Polchinski (1998) for a detailed overview of the Faddeev–Popov gauge fixing for this theory.

The states of the string are related to excitations of certain fields in spacetime. In particular, the Hilbert space $\mathcal{H}_0$ has a basis called the *Fock space basis* which makes clear the relation between worldsheet states and spacetime fields. The basis is formed by acting on the vacuum state $|0\rangle$ with worldsheet creation operators α^μ_{-n}, b_{-n}, and c_{-n}, which are the modes of the fields X^μ, b^{ab}, and c_a; the subscript $-n$ denotes the eigenvalue under the (holomorphic) worldsheet dilatation operator L_0 and is called the (left) *level*. For the remainder of this section, we'll ignore the ghost part of the Hilbert space for simplicity. Note that there are two sets of each of the oscillators: α^μ_{-n} and $\bar{\alpha}^\mu_{-n}$. This is because the closed string theory decouples into left-moving (holomorphic) and right-moving (antiholomorphic) sectors. The sectors are related by the *level-matching condition*, which says that the left and right levels of states must be the same.[7,8] The vacuum state is degenerate due to the spacetime translation invariance, and the other vacua are 'spacetime boosts' of $|0\rangle$:

$$|k\rangle = e^{ik_\mu X^\mu} |0\rangle . \tag{10}$$

The Fock space basis is an infinite-dimensional basis given by states of the form

$$\alpha^{\mu_1}_{-n_1} \dots \alpha^{\mu_m}_{-n_m} \bar{\alpha}^{\mu_{m+1}}_{-\bar{n}_1} \dots \bar{\alpha}^{\mu_{m+l}}_{-\bar{n}_l} |k\rangle , \tag{11}$$

where $\sum n_i = \sum \bar{n}_i$ due to the level-matching condition.

The basis state $\alpha^{\mu_1}_{-n_1} \dots \bar{\alpha}^{\mu_N}_{-\bar{n}_N} |k\rangle$ (note the slightly different notation from above) looks like a rank N spacetime tensor, since it has N Lorentz indices. It will typically decompose into different irreducible representations of the spacetime Lorentz group $SO(25,1)$, which act as a global symmetry on the worldsheet CFT:

$$\alpha^{\mu_1}_{-n_1} \dots \bar{\alpha}^{\mu_N}_{-\bar{n}_N} |k\rangle = \bigoplus_{\text{irreps}} (\alpha^{\mu_1}_{-n_1} \dots \bar{\alpha}^{\mu_N}_{-\bar{n}_N})_{\text{irrep}} |k\rangle . \tag{12}$$

For instance, the two-index state $\alpha^\mu_{-1}\bar{\alpha}^\nu_{-1} |0\rangle$ decomposes into a traceless symmetric, antisymmetric, and singlet representation. The component of a general state $|\psi\rangle$ along the basis vector $(\alpha^{\mu_1}_{-n_1} \dots \bar{\alpha}^{\mu_k}_{-n_k})_{\text{irrep}} |0\rangle$ is interpreted as an excitation of a spacetime field which transforms in that representation.[9] For the previous example, the traceless symmetric component of $|\psi\rangle = \zeta_{\mu\nu}\alpha^\mu_{-1}\bar{\alpha}^\nu_{-1} |0\rangle$ is $\frac{1}{2}(\zeta_{\mu\nu} + \zeta_{\nu\mu}) - \zeta^\sigma_\sigma \eta_{\mu\nu} \equiv G_{\mu\nu}$ (the graviton), the antisymmetric component is $\frac{1}{2}(\zeta_{\mu\nu} - \zeta_{\nu\mu}) \equiv B_{\mu\nu}$ (the Kalb–Ramond field), and the singlet component is $\zeta^\sigma_\sigma \eta_{\mu\nu} \equiv \Phi\eta_{\mu\nu}$ (the dilaton). A general state can then be written in the Fock

[7] In terms of operators, $(\bar{L}_0 - L_0) |\psi\rangle \equiv L_0^- |\psi\rangle = 0$. There is also a ghost constraint $(\bar{b}_0 - b_0) |\psi\rangle \equiv b_0^- |\psi\rangle = 0$.

[8] Open strings are given by Riemann surfaces with boundary, and the boundary conditions couple the two sectors completely, meaning there is only one sector. Thus, there is no level-matching condition, and one can have states like $A_\mu(k) e^{ik_\nu X^\nu} \alpha^\mu_{-1} |0\rangle$, which give gauge fields in spacetime.

[9] In the more general case where some other symmetry group G acts on the Hilbert space, we would also distinguish Fock space states based on their representations under that group.

space basis as

$$|\psi\rangle = \int \frac{d^{26}k}{(2\pi)^{26}} \left[T(k) + (G_{\mu\nu}(k) + B_{\mu\nu}(k) + \Phi(k)\eta_{\mu\nu})\alpha^{\mu}_{-1}\bar{\alpha}^{\nu}_{-1} + \dots\right] |k\rangle \,. \tag{13}$$

The components in this basis $T(k), G_{\mu\nu}(k), \dots$ are interpreted as fields in momentum space. In position space, the graviton, for example, is given by

$$G_{\mu\nu}(x) = \int \frac{d^{26}k}{(2\pi)^{26}} e^{ikx} G_{\mu\nu}(k), \tag{14}$$

$$g_{\mu\nu} = \eta_{\mu\nu} + G_{\mu\nu}(x). \tag{15}$$

The last line is meant to illustrate that the spacetime metric *in this state* is not η, but rather $\eta + G$. This is analogous to spin-2 theories of gravity, as will be discussed in §4. Thus, the states of the worldsheet theory are related to excitations of fields in spacetime.

In this basis, it is easier to see what the gauge equivalence (9) means. The operator Q_B gives rise to factors of k_μ when acting on states, and in position space this turns into a derivative ∂_μ. Then, for example, the graviton component $G_{\mu\nu}$ inherits a spacetime gauge invariance[10]

$$G_{\mu\nu} \sim G_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu, \tag{16}$$

which is an infinitesimal spacetime diffeomorphism generated by the vector field ξ_μ. Thus, the gauge invariance (9) gives rise to all gauge invariances of all spacetime fields. The condition $Q_B |\psi\rangle = 0$ gives the mass-shell condition in spacetime, $k^2 = -M^2$ for the various fields. The mass M is related to the level by $M^2 = \frac{1}{l_s^2}(n-1)$, where l_s is the string length; thus, states of higher level have higher mass. Since the graviton has $n = 1$, it is massless. Unfortunately, in the bosonic theory, the $n = 0$ state given by the field T is tachyonic: $M^2 = -\frac{1}{l_s^2}$. This suggests that the theory is non-perturbatively unstable. Some implications of tachyons in the spectrum will be discussed in §5, though this will be for *open* string tachyons, which are much better understood.

2.2.2 Interactions

The *operator-state correspondence* is a special relation in 2D CFT between the Hilbert space of the theory and the set of operators acting on that Hilbert space. In effect, for every state, we have a local operator called the *vertex operator*, and vice versa. If we wish to calculate a scattering amplitude between two asymptotic states $|\psi\rangle$ and $|\phi\rangle$, we can instead find the associated vertex operators O_ψ and O_ϕ and compute the correlation function $\langle O_\psi O_\phi \rangle$. When the state is a Fock space basis vector $(\alpha^{\mu_1}_{-n_1} \dots \bar{\alpha}^{\mu_i}_{-n_i})_{\text{irrep}} e^{ik_\mu X^\mu} |0\rangle$, the associated vertex operator will be denoted $V_j(k, \sigma)$, where k is the momentum of the state, j

[10] One-index spacetime states in the closed string theory arise from ghost factors in the gauge parameter state $|\lambda\rangle$, which we suppressed in the above Fock space expansion.

is a general index specifying the irrep, and σ is the point on the worldsheet where the operator is inserted. This vertex operator can be thought of as producing a momentum k '1-particle' excitation of the field associated to that representation.

The S-matrix between asymptotic states is therefore related to some correlation function of vertex operators. The correlation function is computed using the gauge-fixed Polyakov path integral with local vertex operator insertions, which can be viewed as summing over all intermediate worldsheet geometries with fixed asymptotic states. This is akin to summing over Feynman diagrams with fixed external states. Crucially, in string theory *there is no freedom in choosing the interactions*, unlike in QFT; the Polyakov action already fixes all string interactions.

One must be careful to preserve the gauge symmetry of the path integral. In particular, to preserve 2D diffeomorphism invariance, the vertex operators must be integrated over the entire worldsheet. Furthermore, to preserve Weyl invariance, the states associated to the vertex operators *must be in the BRST cohomology*. Hence, in the worldsheet formulation of string theory, one can only compute interactions between on-shell states with $k^2 = -M^2$, which means that it is difficult to compute off-shell quantities such as potentials for fields. This situation will change once we move to SFT.

There is one further technical interjection that will later prove important in the formulation of SFT. The Polyakov path integral requires an integration over the worldsheet metric h_{ab} on the Riemann surface Σ_g. Due to the symmetries (2) and properties of 2D metrics, this infinite-dimensional integral is *almost* trivial. That said, there are some global technicalities due to 'modular parameters', which are deformations of the metric that are neither diffeomorphisms nor Weyl transformations. These parameters can be thought of as a 'change of shape' of the Riemann surface; for example, on the torus $T^2 = S^1 \times S^1$, these transformations roughly change the relative size/rotation of the two circles. The number of modular parameters is 0 for the sphere, 2 for the torus, and $6g - 6$ for a Riemann surface of genus $g \geq 2$. We can introduce a 0, 2, or $6g - 6$ dimensional space $\mathcal{M}_g$ called the *moduli space of Riemann surfaces of genus g*. This space is a manifold with singularities (i.e., an orbifold), and points in this space represent Riemann surfaces with particular values for the modular parameters. The path integral over metrics then reduces to an integral over modular parameters, which is an integral over this *finite* dimensional space. There is a measure over this space defined in terms of the b-ghost, and the path integral naturally accounts for this.[11] There is also a subtlety due to Diff $\times$ Weyl transformations that leave the metric invariant. These are called *conformal Killing vectors* (CKVs); the sphere has 6, the torus has 2, and higher genus Riemann surfaces have 0. CKVs can be accounted for by inserting factors of c into correlation functions.

In its full glory, the bosonic string S-matrix for external states (i.e. vertex operators) V_i with momenta k_i with the subtleties of the modular parameters

[11] See e.g. Nakahara (2003).

(and CKVs) accounted for, is

$$S_{j_1\dots j_n}(k_1,\dots,k_n) = \sum_g \int_{\mathcal{M}_g} \frac{d^{\dim(\mathcal{M}_g)}t}{n_R} \int \mathcal{D}X\mathcal{D}b\mathcal{D}c\, e^{-S_P - S_{\rm gh} - \lambda\chi} \times \prod_{(a,i)\notin f} \int d\sigma_i^a \prod_{k=1}^{\dim(\mathcal{M}_g)} \frac{1}{4\pi}(b, \partial_k \hat{h}(t)) \prod_{(a,i)\in f} c^a(\hat{\sigma}_i) \prod_{i=1}^{n} \hat{h}(\sigma_i,t)^{1/2} V_{j_i}(k_i,\sigma_i). \tag{17}$$

The details needn't detain us. The rough form is just a correlation function of the vertex operators with various ghost insertions

$$S \sim \sum_g e^{-\lambda\chi} \int_{\mathcal{M}_g} d^{\dim(\mathcal{M}_g)}t \int_{\Sigma_g} d^2\sigma\, \langle (b,\partial_t h)\dots c\dots V_1\dots V_n\rangle_{\Sigma_g}\,. \tag{18}$$

The vertex operators are integrated over the Riemann surface, the correlation function is integrated over all the moduli, and there is a sum over the genus, which can be interpreted as a loop expansion as in QFT. The different genus contributions are weighted by the Euler characteristic $\chi = 2-2g$; thus, the quantity e^λ acts like a coupling constant, and will be denoted g_s. This term arises because in 2D, the worldsheet Einstein–Hilbert action $S = \frac{\lambda}{4\pi}\int d^2z\sqrt{h}\mathcal{R}$ is also Diff $\times$ Weyl invariant, so we should also include it in our theory in addition to the Polyakov action. This term is topological due to the Gauss–Bonnet theorem, and it is proportional to the Euler characteristic χ. It has no effect other than weighting the different genus terms with different powers of g_s. As we can see from (5), this corresponds to a constant background dilaton field $\Phi_0 = \lambda \to g_s = e^{\Phi_0}$. Thus, in string theory, the coupling is determined *dynamically* by the value of the dilaton. It is not a free parameter.

2.2.3 Backgrounds in worldsheet string theory

Much of the above discussion was in the context of the Polyakov action (1) formulated around flat, 26D Minkowski space. Suppose we instead want to study bosonic string theory in a spacetime with many background fields of all spins, including a background metric field. In other words, suppose that we want to consider dynamical possibilities with objects $\langle M, \hat{T}, \hat{g}_{\mu\nu}, \hat{B}_{\mu\nu}, \hat{\Phi}, \dots\rangle$.[12] This can be done in two different ways:

1. We can write down a new worldsheet action as in §2.1

$$S'_P[X^\mu, h_{ab}|\hat{T},\hat{g},\hat{B},\hat{\Phi},\dots] := \frac{1}{4\pi}\int_{\Sigma_g} d^2\sigma\sqrt{h}\Big[\hat{T}(X) + \mathcal{R}\hat{\Phi}(X) + \Big(\hat{g}_{\mu\nu}(X)h^{ab} + \hat{B}_{\mu\nu}(X)\epsilon^{ab}\Big)\partial_a X^\mu \partial_b X^\nu + \dots\Big]. \tag{19}$$

[12] From this point onwards, we'll denote background fields with hats; to be clear, this should not be taken to imply that there is anything quantum mechanical about these objects.

This theory can be quantized via a path integral. There will be a Weyl anomaly unless the couplings $\hat{T}, \hat{g}_{\mu\nu}, \hat{B}_{\mu\nu}, \hat{\Phi}, \ldots$ satisfy certain field equations, including, most notably, the Einstein field equations.[13]

2. Instead, we might notice that inserting an operator

$$\exp\left(-\frac{1}{4\pi}\int d^2\sigma\sqrt{h}G_{\mu\nu}(X)h^{ab}\partial_a X^\mu \partial_b X^\nu\right) \tag{20}$$

into the flat space Polyakov path integral has the effect of shifting the metric $\eta_{\mu\nu} \to \eta_{\mu\nu}+G_{\mu\nu}$. Note that this operator is $\exp(-\frac{1}{4\pi g_s}\int d^2\sigma\sqrt{h}e^{-ikX}V_{\text{grav}})$, i.e. it is the exponential of the graviton vertex operator.[14] More generally, if we want a background value for a field corresponding to the vertex operator V_j, we simply insert the operator $\exp(-\frac{1}{4\pi}\int d^2\sigma\sqrt{h}V_j)$ into the path integral; this is a so-called coherent state.[15] Again, this can only be done for BRST invariant operators V_j, which means this is a 'marginal deformation', i.e. preserves conformal invariance. Thus, backgrounds obtained in this way also satisfy the field equations from above.

2.2.4 Superstring theory

In superstring theory, in addition to worldsheet conformal symmetry, we require worldsheet supersymmetry, which is a symmetry that takes bosons into fermions and vice versa. Such theories are called *superconformal field theories* or SCFTs. We can impliment this symmetry by introducing worldsheet fermions into the action (written in complex worldsheet coordinates $z = \sigma + i\tau$):

$$S = \frac{1}{4\pi}\int d^2z\left(\partial X^\mu\bar{\partial}X_\mu + \psi^\mu\bar{\partial}\psi_\mu + \tilde{\psi}^\mu\partial\tilde{\psi}_\mu\right) + S_{\text{ghost}}. \tag{21}$$

Fermions can be either periodic or antiperiodic when going around the string. Periodic fermions are said to be in the R (Ramond) sector, while antiperiodic fermions are in the NS (Neveu–Schwarz) sector. For the anomaly to vanish in this theory, we find $D = 10$.

The action (21) is for the 'type II' closed string theory, which has a holomorphic SCFT and an antiholomorphic SCFT.[16] It has four sectors depending on the periodicities of ψ and $\tilde{\psi}$: NSNS, RNS, NSR, and RR. The NSNS states are familiar from bosonic string theory and include the graviton, Kalb–Ramond B-field, and dilaton. The RNS and NSR sectors are fermions *in spacetime* and in fact give rise to an $\mathcal{N} = 2$ supersymmetry in spacetime.[17] The RR sector

[13] See Huggett and Vistarini (2015) and Read (2019) for philosophical discussion.

[14] Sometimes, the integral over the worldsheet $\int d^2\sigma\sqrt{h}$ is incorporated into the definition of the vertex operator.

[15] For philosophical discussion, see Huggett and Wüthrich (2025) and Read (2023).

[16] The distinction between IIA and IIB superstring theories arises from a difference in *GSO projection* between the two sectors, which truncates the spectrum and is needed to have a consistent theory.

[17] *A priori* supersymmetry on the worldsheet and supersymmetry in spacetime have nothing to do with each other, so this is a surprising fact from the worldsheet point of view.

has new bosonic higher-form field strengths, and these pose certain difficulties which will be discussed later.

The heterotic strings are formed by combining a holomorphic SCFT with an antiholomorphic bosonic CFT.[18] Thus we just have an R sector and an NS sector depending on the periodicity of ψ, which are fermions and bosons in spacetime respectively. There is an $\mathcal{N} = 1$ supersymmetry in spacetime, as well as a spacetime gauge group SO(32) or $E_8 \times E_8$.

In addition to the diffeomorphisms and Weyl transformations, we need to gauge fix the local supersymmetry. This is done by introducing additional ghost fields, which again have opposite statistics to the fields to which they are associated. Since supersymmetries are generated by anti-commuting spinors, the ghosts will be commuting fields. They are denoted as the '$\beta\gamma$' CFT; the holomorphic part of the ghost action is

$$S_g = \frac{1}{2\pi} \int d^2z (b\bar{\partial}c + \beta\bar{\partial}\gamma). \tag{22}$$

Recall that if we wish to compute scattering amplitudes between states, we must find the appropriate vertex operators corresponding to those states. In the superstring case, there turn out to be difficulties in this process due to the definition of the vacuum. Essentially, one finds that the operator corresponding to the vacuum must satisfy very strange relations from the perspective of the $\beta\gamma$ CFT. For instance, the NS vacuum $|0\rangle_{\text{NS}}$ maps to the delta function $\delta(\gamma)$!

Thus, instead of using the $\beta\gamma$ ghosts, it is useful to introduce an equivalent CFT formulated in terms of an anticommuting ghost system '$\xi\eta$' (which are like the bc ghosts) and a free boson ϕ:

$$\beta(z) \cong e^{-\phi(z)}\partial\xi(z), \ \gamma \cong e^{\phi(z)}\eta(z). \tag{23}$$

In terms of these variables, the vacua are easily expressed as $|0\rangle_{\text{NS}} \cong e^{-\phi}$ and $|0\rangle_{\text{R}} \cong e^{-\phi/2}$. Under the current $j = \partial\phi$, these operators have charge -1 and $-1/2$ respectively; this charge is called the *picture number*. Since general physical states are constructed by acting on the vacuum with matter creation operators, general vertex operators in the NS and R sectors will also have picture number -1 and $-1/2$ respectively, i.e., they will contain factors of $e^{-\phi}$ and $e^{-\phi/2}$ respectively.

One must also worry about the fermionic analogues of the moduli and CKVs when computing superstring scattering amplitudes. Recall that in the bosonic case, we needed to insert factors of b and c to compensate for these subtelties. In the fermionic case, we must instead insert *picture changing operators* (PCOs):

$$\mathcal{X}(z) := \{Q_B, \xi\}. \tag{24}$$

[18] There is a mismatch in central charge ($c = 15$ for the holomorphic part and $\bar{c} = 26$ for the antiholomorphic part) between the two sectors; on the bosonic side, the interpretation of a particular boson as a coordinate only holds for 10 bosons, so $D = 10$ in spacetime. The remaining 16 bosons are interpreted as giving rise to the gauge group. For philosophical discussion of this, see Krátký and Read (2024).

The explicit form of the PCO is quite messy, but it is crucial in superstring scattering amplitudes and shows up in the formulation of superstring field theory. We will just use the zero mode of this operator

$$\mathcal{X}_0 = \frac{1}{2\pi i} \oint \frac{1}{z} \mathcal{X}(z). \tag{25}$$

As the name suggests, $\mathcal{X}_0$ adds 1 to the picture number of the amplitude.

The superstring S-matrix has a very similar form to the bosonic case but with additional $\beta\gamma$ and $\mathcal{X}_0$ insertions. Background fields are much more subtle in the superstring case, however (see e.g. Cho et al. (2020)). Background RR sector fields in particular *cannot* be incorporated using a straightforward worldsheet action or the coherent state method; roughly, the reason is because R sector vertex operators are worldsheet spinors, and they cannot be exponentiated in the right way. String field theory will provide a workaround for this difficulty.

3 String field theory

Perturbative string theory based upon the Polyakov action has its shortcomings. In particular, there are issues with IR divergences and renormalization,[19] as well as the aforementioned problem with background RR fields. For these reasons and others, it is therefore useful to consider a second-quantized version of the string, i.e. a 'string field theory' (SFT).[20] Roughly, the SFT is a special QFT whose scattering amplitudes reproduce all Polyakov scattering amplitudes (17). This might seem contradictory: the Polyakov scattering amplitudes include UV *finite* graviton scattering amplitudes, and a standard QFT would not be able to accommodate this due to the nonrenormalizability of gravity in higher dimensions. Indeed, SFT gets around this issue with various subtleties: it has infinite fields of higher mass and spin which 'tame' the UV behaviour, the action is in general nonlocal in the position space representation, and it has an intricate gauge structure.

SFTs are based on the string field $|\Psi\rangle \in \mathcal{H}$, which is an arbitrary (Grassman even) member of the Hilbert space of the worldsheet theory satisfying the level matching condition $L_0^- |\Psi\rangle = 0$.[21] It might look odd having the field written as a ket instead of an operator, but it can be translated into a more conventional field by using the Schrödinger representation, which is a functional in spacetime $\langle X^\mu, c|\Psi\rangle = \Psi[X^\mu(\sigma), c]$. The string field depends on the shape of the string in spacetime (and also the c ghost), not just a point like a conventional quantum field; this is the source of nonlocality in the interactions. We will also use the symbol Ψ to represent the vertex operator associated to $|\Psi\rangle$.

Although the string field can be represented as a functional on loop space $\Psi[X^\mu(\sigma), c]$, if we use the Fock space basis of the Hilbert space, we can expand it

[19] There are no UV divergences due to the properties of moduli spaces of Riemann surfaces that control loop integrals.

[20] For recent reviews of SFT, see Erbin (2021) and Sen and Zwiebach (2024).

[21] This is needed for consistent interactions. Henceforth, $\mathcal{H} \subset \mathcal{H}_0$ will denote the subspace of the Hilbert space which meets these conditions.

as a collection of infinite *point-valued* fields, which represent the various modes of the string. One defines the quantum theory by doing a path integral over Ψ, which amounts to doing infinite path integrals over each of the components of Ψ, i.e., infinite path integrals over these classical fields.

3.1 Bosonic closed string field theory

The kinetic part of the action for the bosonic closed string field theory has the form[22]

$$\frac{1}{2}\langle\Psi|\, c_0^- Q_B\,|\Psi\rangle\,, \tag{26}$$

which tells us that on-shell states are in the BRST cohomology,

$$Q_B\,|\Psi\rangle = 0. \tag{27}$$

This is analogous to the Klein-Gordon part of scalar field theory: $(\Box - m^2)\phi = 0$, which tells us that on-shell states satisfy $-k^2 = m^2$, exactly like the BRST condition. The free theory only has this kinetic term in its action, and so it has a gauge symmetry given by $|\Psi\rangle \to |\Psi\rangle + Q_B\,|\Lambda\rangle$, since $Q_B^2 = 0$.

To get a better picture of what this means, we can write down a spacetime action by expanding the string field in the Fock space basis of the Hilbert space (where we ignore ghost factors for simplicity):

$$|\Psi\rangle = \int \frac{d^{26}k}{(2\pi)^{26}}[T(k) + (G_{\mu\nu}(k) + B_{\mu\nu}(k) + \Phi(k)\eta_{\mu\nu})\alpha^\mu_{-1}\bar{\alpha}^\nu_{-1} + \dots]\,|k\rangle\,. \tag{28}$$

The coefficients $T(k), G(k), \dots$ are interpreted as Fourier transforms of the spacetime fields $T(x), G(x), \dots$. In position space, the action (26) gives kinetic terms for all of the fields:

$$S = \int d^{26}x \left(\frac{1}{2}\eta^{\mu\nu}\partial_\mu T(x)\partial_\nu T(x) + \dots\right), \tag{29}$$

where we have assumed our worldsheet theory is 26D flat Minkowski space. The free SFT is therefore a quantum field theory with infinite fields which represent the various vibrational modes of the strings. This move to position space will not work for the action of the interacting theory.

To write down the interacting SFT action, we need to go beyond the Polyakov S-matrix from above. We need amplitudes of off-shell vertex operators ($Q_B\,|V\rangle \neq 0$), which the Polyakov formalism cannot deal with (see Polchinski (1998, p. 103)). Once these amplitudes are known, we can reverse-engineer an action whose 'Feynman diagram' vertices give rise to them.

It is therefore necessary to consider off-shell string states which are not necessarily conformally invariant, which means they will be coordinate-dependent.[23]

[22] The $c_0^- = c_0 - \bar{c}_0$ ghost is part of the definition of the inner product.

[23] This is because we need to insert factors of $\int d^2z V$ into amplitudes, and such factors will be coordinate dependent unless V is a $(1,1)$ primary (i.e. $Q_B\,|V\rangle = 0$), since the transformation of d^2z will not cancel that of V.

This can be done by introducing local coordinates around each puncture on the Riemann surfaces where the vertex operators are inserted. We can systematically account for this using a fiber bundle with base space $\mathcal{M}_{g,n}$, the moduli space of Riemann surfaces of genus g with n punctures,[24] and with the fiber being a choice of local coordinates around each puncture (up to a phase). This bundle is denoted $\hat{\mathcal{P}}_{g,n}$.[25]

We now want to look for an amplitude constructed with this space which reproduces Polyakov amplitudes when the vertex operators are on-shell but which still works off-shell. We begin by simply picking specific local coordinates for each Riemann surface with fixed moduli and vertex operator positions, i.e. we pick a specific section $\mathcal{F}_{g,n}$ of $\hat{\mathcal{P}}_{g,n}$. If we integrate over $\mathcal{F}_{g,n}$, then we will also be integrating over all of the moduli and all of the vertex operator positions, since both are already accounted for in the base space $\mathcal{M}_{g,n}$. Thus, to get the genus g contribution to the n-point amplitude, we need to integrate some $p = \dim(\mathcal{M}_{g,n})$ form over this section:

$$\mathcal{A}_g(V_1, \ldots, V_n) = (g_s)^{-\chi_{g,n}} \int_{\mathcal{F}_{g,n}} \Omega_p^{(g,n)}(V_1, \ldots, V_n). \tag{30}$$

The factor of g_s plays the role of the coupling constant and appears for the same reason as in (17). The specific p-form can be determined by the condition that it reproduces (17) when the V_i are all on-shell:

$$\Omega_p^{(g,n)}\left[\frac{\partial}{\partial u^{j_1}}, \ldots, \frac{\partial}{\partial u^{j_p}}\right] = \left(-\frac{1}{2\pi i}\right)^{3g-3+n} \left\langle \mathcal{B}\left[\frac{\partial}{\partial u^{j_1}}\right] \ldots \mathcal{B}\left[\frac{\partial}{\partial u^{j_p}}\right] V_1 \ldots V_n \right\rangle. \tag{31}$$

$\mathcal{B}\left[\frac{\partial}{\partial u}\right]$ is an operator constructed from the b-ghosts and is a one-form on the tangent space of $\hat{\mathcal{P}}_{g,n}$. The total amplitude is then

$$\mathcal{A}(V_1, \ldots, V_n) = \sum_{g=0}^{\infty} \mathcal{A}_g(V_1, \ldots, V_n). \tag{32}$$

Essentially, the amplitude has the form of a normal Polyakov string amplitude (18) with vertex operators and ghost insertions, just with the new detail of local coordinates, which modify the ghost insertions slightly and require the integration over the section $\mathcal{F}_{g,n}$ instead of just $\mathcal{M}_{g,n}$. The integration over $\mathcal{F}_{g,n}$ contains the integration over moduli *and* vertex operator positions, so it combines the two types of integrals we saw in the S-matrix. Note that all physical quantities are independent of the choice of section, which means they are independent of the local coordinates.

[24] This space is $\dim(\mathcal{M}_g) + 2n$ dimensional, where the dimension of $\mathcal{M}_g$ is discussed above. This is because this space includes both the geometric moduli as well as the 2 coordinates of each of the n punctures.

[25] A point in this bundle will be a Riemann surface of genus g with n punctures with a choice of local coordinates around each of these punctures. The projection $\pi : \hat{P}_{g,n} \to \mathcal{M}_{g,n}$ simply 'forgets' the local coordinates. Note that the fiber itself is infinite dimensional.

We now want to reverse-engineer an action by using the amplitudes. However, we need to be careful. The difficulties are illustrated in a simple example from QFT (drawn from Erbin (2021, p. 295)). Suppose we have a $\phi^3 + \phi^4$ scalar field theory. Now consider a $2 \to 2$ scattering. There are tree-level contributions from the ϕ^3 term, with two interactions connected by an internal propagator. However, since there is also a ϕ^4 interaction, there will also be a tree-level interaction with all particles connecting at one point. If we are trying to reverse-engineer the action by just looking at amplitudes, we need to properly account for both of these effects. For example, if we know all of the lower-order interactions up to the n-point scattering (here the ϕ^3 term), and we know the propagator, then we can deduce the existence of a ϕ^4 term in the action, since the ϕ^3 diagrams alone do not give the full amplitude.

Similarly, in SFT, the n-point amplitude at genus g (30) will have many contributions coming from other-order interactions with propagators. However, some parts of the amplitude will not be accounted for if we just use these interactions, and so we must add a new n-point interaction term in the action. This will continue for all n, so there will be infinite interaction terms.[26] To describe these terms explicitly, we need to understand how the other-order interactions contribute to the n-point amplitudes.

There are two important gluing operations on the moduli spaces of Riemann surfaces that allow one to construct higher genus Riemann surfaces from lower genus Riemann surfaces. In particular, starting with two punctured Riemann surfaces $\Sigma^1_{g_1,n_1}$ and $\Sigma^2_{g_2,n_2}$, we can construct a new Riemann surface $\Sigma_{g,n}$ of genus $g = g_1 + g_2$ with $n = n_1 + n_2 - 2$ punctures by gluing a puncture of Σ^1 with a puncture of Σ^2. This can be denoted by $\{\Sigma^1, \Sigma^2\} = \Sigma$. The other gluing operation just glues two punctures of the same Riemann surface $\Sigma^3_{g_3,n_3}$ which gives a new Riemann surface $\Sigma_{g,n}$ of genus $g = g_3 + 1$ and $n = n_3 - 2$ punctures. This operation can be denoted $\Delta\Sigma^3 = \Sigma$.[27]

These two operations can be interpreted as follows. The operation of gluing two distinct Riemann surfaces should be understood as two separate Feynman diagrams with a propagator running between them. In other words, this is a 1PR (one particle reducible) graph with all interactions coming from lower-order vertices. Similarly, the operation of gluing a surface to itself gives the contribution coming from higher point vertices with internal loops. This latter effect did not show up in the tree-level $\phi^3 + \phi^4$ example from above, since it arises from loops.

Now, by the logic above, in order to find the n-point interaction term in the action, we need to look at the part of (30) that is not accounted for by other-order interactions that are glued together. If we glue all possible Riemann surfaces of appropriate g_i and n_i in all possible ways, we will cover a region of $\mathcal{M}_{g,n}$.[28] However, there will be some subset of $\mathcal{M}_{g,n}$ that is not covered. This

[26] An important exception is Witten's cubic open string field theory, which can reproduce all interactions with just a cubic term in the action. This will be discussed in §3.6.

[27] In both of these operations, there is an arbitrary complex parameter which determines exactly how the gluing is performed. This is discussed in more detail in the next footnote.

[28] We might worry about covering the same region twice, which would over-count certain

in turn will correspond to some subset of the section $\mathcal{F}_{g,n}$; it is denoted $\mathcal{V}_{g,n}$ and is called a *string vertex*. In the language of the QFT analogy, this string vertex is akin to the ϕ^4 interaction. It must appear in the action because the other interaction terms are not able to reproduce the full amplitude by themselves. As one might expect, these vertices must satisfy complicated and strict geometric relations so as to perfectly cover all moduli spaces of Riemann surfaces. This will give rise to a geometric BV master equation, which is discussed in §3.3.

Using the string vertices $\mathcal{V}_{g,n}$, we can define a multilinear bracket $\{A_1, \ldots, A_n\} : \mathcal{H}^{\otimes n} \to \mathbb{C}$,

$$\{A_1, \ldots, A_n\} := \sum_{g=0}^{\infty} (g_s)^{-\chi_{g,n}} \int_{\mathcal{V}_{g,n}} \Omega^{(g,n)}_{\dim(\mathcal{M}_{g,n})}(A_1, \ldots, A_n). \tag{33}$$

The action of the interacting quantum bosonic closed string field theory is then

$$S = \frac{1}{2} \langle \Psi | \, c_0^- Q_B \, | \Psi \rangle + \sum_{n=1}^{\infty} \frac{1}{n!} \{\Psi^n\}, \tag{34}$$

where

$$\{\Psi^n\} := \underbrace{\{\Psi, \ldots, \Psi\}}_{n}. \tag{35}$$

Due to the geometric properties of the vertices, this action will be able to reproduce all bosonic closed string Polyakov amplitudes.

This action has a complicated gauge invariance:

$$|\Psi\rangle \to |\Psi\rangle + Q_B \, |\Lambda\rangle + \sum_{n=1}^{\infty} \frac{1}{n!} \, |[\Lambda \Psi^n]\rangle \, , \tag{36}$$

where

$$\langle A_0 | \, c_0^- \, |[A_1, \ldots, A_n]\rangle := \{A_0, \ldots, A_n\} \tag{37}$$

defines the *string bracket*: $|[A_1, \ldots, A_n]\rangle : \mathcal{H}^{\otimes n} \to \mathcal{H}$. A common choice to gauge fix this symmetry is $b_0^- \, |\Psi\rangle = 0$, which is called *Siegel gauge*. The classical equation of motion is

$$Q_B \, |\Psi\rangle + \sum_{n=1}^{\infty} \frac{1}{n!} \, |[\Psi^n]\rangle = 0. \tag{38}$$

Quantum observables will be discussed in §3.3, since the gauge invariance (36) will restrict the space of physical operators.

---

contributions. This can be avoided by adjusting a parameter in the gluing operation called the *stub parameter*. Although the stub parameter can change the form of the string vertices defined below, it can be shown that SFTs with different stub parameters are related by field redefinitions and are hence equivalent.

3.2 Closed superstring field theory

Superstring field theory is structurally very similar to bosonic string field theory. One again has a physical string field Ψ, now defined to have picture number -1 in the NS sector and $-1/2$ in the R sector. The reason for these assignments is the same as discussed in §2.2.4.

A major difference in the superstring case, however, is the introduction of an auxiliary, unphysical string field $\tilde{\Psi}$, which is defined to have picture number -1 in the NS sector and $-3/2$ in the R sector. This field is needed to write down a kinetic term for the R sector. As a motivation for its introduction, one might recall the presence of the self-dual RR 5-form in type IIB supergravity/string theory. The difficulties in writing an action for this field are well known. Since this field should arise within the type II superstring field theory, similar issues arise here. The auxiliary field, however, provides a workaround for this difficulty.[29]

Define $\mathcal{G}$ to be 1 in the NS sector and $\mathcal{X}_0$, the zero mode of the picture changing operator (25), in the R sector. Then the closed superstring field theory action is

$$S = -\frac{1}{2}\langle\tilde{\Psi}|\, c_0^- \mathcal{G} Q_B\,|\tilde{\Psi}\rangle + \langle\tilde{\Psi}|\, c_0^- Q_B\,|\Psi\rangle + \sum_{n=1}^{\infty}\frac{1}{n!}\{\Psi^n\}, \tag{39}$$

and the equations of motion become

$$Q_B\,|\Psi\rangle - \mathcal{G} Q_B\,|\tilde{\Psi}\rangle = 0 \tag{40}$$

$$Q_B\,|\tilde{\Psi}\rangle + \sum_{n=1}^{\infty}\frac{1}{n!}\,|[\Psi^n]\rangle = 0, \tag{41}$$

which simplify to

$$Q_B\,|\Psi\rangle + \sum_{n=1}^{\infty}\frac{1}{n!}\mathcal{G}\,|[\Psi^n]\rangle = 0. \tag{42}$$

$$Q_B\,|\Psi\rangle = \mathcal{G} Q_B\,|\tilde{\Psi}\rangle\,. \tag{43}$$

One can solve these equations by finding Ψ using the first equation and $\tilde{\Psi}$ with the second. We will not write $\tilde{\Psi}$ explicitly when describing solutions of the superstring field theory, since it is completely determined by Ψ. This action

[29] Another difference from the bosonic case is that the section $\mathcal{F}$ has to be replaced with a more general object called a *chain*, which must appropriately project down onto $\mathcal{M}$. The reason is that picture changing operators can collide with each other and with vertex operators, which lead to *spurious divergences*. A technique called *vertical integration* is needed to avoid these collisions, and it requires vertical segments on $\mathcal{F}$, which means it is no longer a section.

also has a gauge invariance:

$$|\Psi\rangle \to |\Psi\rangle + Q_B |\Lambda\rangle + \sum_{n=2}^{\infty} \frac{1}{n!} \mathcal{G} |[\Lambda \Psi^n]\rangle \tag{44}$$

$$|\tilde{\Psi}\rangle \to |\tilde{\Psi}\rangle + Q_B |\tilde{\Lambda}\rangle + \sum_{n=2}^{\infty} \frac{1}{n!} |[\Lambda \Psi^n]\rangle \,. \tag{45}$$

3.3 BV quantization and L_∞ algebras

The construction of the SFT action might well have seemed rather *ad hoc* until this point. There were many free choices in the procedure, including the precise choice of vertices and the choice of section over the bundle. There is a much more abstract perspective on the action which makes it clear that these choices are irrelevant, and one can explicitly prove that the various SFTs one obtains from these constructions are equivalent.

When we decomposed the moduli spaces of Riemann surfaces into various vertices, we needed to ensure that these vertices perfectly covered the moduli spaces without any gaps or overlaps. This consistency condition can be expressed as a *geometric BV master equation*:

$$-\partial \mathcal{V}_{g,n} = \Delta \mathcal{V}_{g-1,n+2} + \frac{1}{2} \sum_{\substack{g_1+g_2=g \\ n_1+n_2-2=n}} \{\mathcal{V}_{g_1,n_1}, \mathcal{V}_{g_2,n_2}\}. \tag{46}$$

To explain the notation, recall that $\mathcal{V}_{g,n}$ is a section over a subspace of $\mathcal{M}_{g,n}$. $\partial \mathcal{V}_{g,n}$ is defined as the boundary of the projection of this subspace down to $\mathcal{M}_{g,n}$, i.e. the boundary of $\pi(\mathcal{V}_{g,n})$. The operations $\Delta\mathcal{V}$ and $\{\mathcal{V}_1, \mathcal{V}_2\}$ are defined as the spaces formed by performing the operations $\Delta\Sigma$ and $\{\Sigma_1, \Sigma_2\}$ for all $\Sigma \in \pi(\mathcal{V})$ and $\Sigma_1 \in \pi(\mathcal{V}_1), \Sigma_2 \in \pi(\mathcal{V}_2)$ respectively. In other words, similarly to above, we glue all Riemann surfaces in the vertices of appropriate g, n using the operations ΔA and $\{A, B\}$. Addition means union, and the negative indicates that the orientations are opposite.

The above equation is so-named because of its algebraic relation to another equation called the *BV master equation*:

$$\Delta S + \frac{1}{2}\{S, S\} = 0. \tag{47}$$

This equation has its origins in the BV quantization of gauge theories, which is the most powerful technique for quantizing theories with intricate gauge symmetries.[30] BV quantization is extremely powerful for understanding string field theory and its gauge-invariant observables, so let us quickly review how it works.

Suppose we want to quantize a classical action S_0 which has some gauge symmetry G, and perhaps even gauge symmetries of gauge symmetries G', and

[30] See e.g. Cattaneo et al. (2025).

so on.[31] One introduces ghosts for each generator of the gauge group (and ghost of ghosts, and so on). Then, the set of physical fields and ghosts is collectively called the set of *fields*, denoted ψ^a. For every field, we now introduce an *antifield* $\tilde{\psi}^a$. Suppose we have N fields, i.e. $a : 1, \ldots, N$. We can then combine the fields and antifields into a $2N$ dimensional (super)manifold M whose coordinates are given by $(\psi^a, \tilde{\psi}^a)$. We will use a combined index $I : 1, \ldots, 2N$ such that for $I : 1, \ldots, N$, $\psi^I = \psi^a$, and for $I : N+1, \ldots, 2N$, $\psi^I = \tilde{\psi}^a$. Introduce a sympectic form on this space called ω^{IJ}.[32] Using this form, we can define the BV *antibracket* between functions on M:

$$\{F, G\} = \frac{\partial_r F}{\partial \psi^I} \omega^{IJ} \frac{\partial_l G}{\partial \psi^J}, \tag{48}$$

and we can define the BV *Laplacian*,

$$\Delta F = \frac{1}{2}(-1)^{\psi^I} \frac{\partial_l}{\partial \psi^I} \left(\omega^{IJ} \frac{\partial_r F}{\partial \psi^J} \right). \tag{49}$$

The subscripts r and l and the exponent $(-1)^{\psi^I}$ have to do with the Grassmanality of the (anti)fields, which we will not discuss for simplicity. To quantize the theory S_0, we first extend it to a function S on M that reduces back to S_0 when the antifields are set to zero:

$$S_0(\psi^a) \to S(\psi^a, \tilde{\psi}^a) : S(\psi^a, 0) = S_0(\psi^a). \tag{50}$$

If this extended action S, called the *quantum master action*, satisfies the above BV master equation (47), then it will define an anomaly-free, gauge-invariant quantum theory. Gauge fixing involves choosing an arbitrary (odd) function F on M and imposing the condition $\tilde{\psi}^a = \frac{\partial_l F}{\partial \psi^a}$. Then, the quantum theory is based on the gauge fixed path integral[33]

$$\int \prod\nolimits_a d\psi^a e^S. \tag{51}$$

The gauge-invariant observables A of this quantum theory are functions on M that satisfy

$$\Delta A + \{S, A\} = 0, \tag{52}$$

and their expectation values are determined by

$$\langle A \rangle = \int_L \mathcal{D}\psi^a A(\psi^a) e^S, \tag{53}$$

[31] This happens in higher-form theories. Suppose we have a theory based on an abelian two form potential B whose field strength is $H = dB$. The potential B has a gauge transformation under $B \to B + d\Lambda$, where Λ is a one form. However, Λ also has a gauge transformation under $\Lambda \to \Lambda + d\lambda$, where λ is a zero form.

[32] This is the inverse matrix of the components of the symplectic form ω_{IJ} in the coordinates ψ^I.

[33] Note the unusual sign in the exponent.

where L is a Lagrangian submanifold of the full supermanifold M.[34]

Indeed, the SFT action (34) satisfies the master equation (47): the space of field configurations M is just the Hilbert space $\mathcal{H}$ of the worldsheet theory, the antibracket is determined in terms of the CFT inner product, and the split into fields and antifields is based on the worldsheet ghost number of the state of the CFT. Notice that this is a somewhat backwards story to the one which we told above: there we had a classical action we were trying to quantize, and we had to introduce ghosts, antifields, and other objects. In SFT, the entire BV structure is already there, and we already have a gauge-fixed master action (when we impose Siegel gauge). It follows that SFT is a consistent, anomaly-free, gauge-invariant quantum theory. The quantum observables of string field theory are functions $F(\Psi)$ satisfying

$$\Delta F + \{S, F\} = 0, \tag{54}$$

and their expectation values are given by

$$\langle F \rangle = \int_L \mathcal{D}\Psi F(\Psi) e^{S(\Psi)}. \tag{55}$$

The algebra satisfied by $\{,\}$ and Δ (potentially including more operations given by brackets $[\cdot]$) is called a (generalized) *BV algebra*. Then, in more abstract terms, the SFT is determined by a morphism from the BV algebra of the vertices, which is defined by the decomposition of the moduli spaces of Riemann surfaces, into the BV algebra of multilinear functions on the Hilbert space of the CFT, which is defined by the string brackets. Due to the properties of the BRST charge and its action on vertex operators, these brackets satisfy the axioms of a *quantum* L_∞ *algebra*.[35]

One can show that different SFTs based on the same background, i.e. different SFTs based on the same worldsheet theory with Hilbert space $\mathcal{H}$, are related by L_∞ isomorphisms (Kajiura 2002; Muenster and Sachs 2012). Moreover, such isomorphisms preserve the action of the theory. Thus, there is a good case that L_∞ isomorphism is the correct standard for the equivalence of SFTs. Since all string field theories formulated using the above construction will end up with the same L_∞ structure, but merely with different definitions for the string vertices and sections, we see that there is a unique SFT corresponding to every worldsheet background, and the various choices made in formulating the theory are irrelevant.

Concretely, the different L_∞ isomorphisms are field redefinitions of the string field that preserve the BV structure and action. In §4.2, we will focus on background independence proofs that use this latter (more familiar) notion of equivalence, though the L_∞ approach has also produced results pointing in similar directions (Muenster and Sachs 2012).

[34] A Lagrangian submanifold is a submanifold where the symplectic form vanishes, $\omega = 0$.

[35] See Sen and Zwiebach (2024) for a detailed introduction to L_∞ (and A_∞) structures.

3.4 The 1PI action

Similar to standard QFT, we can write down a 1PI (one particle irreducible) effective action for SFT. Recall that such an action is defined as the Legendre transformation of the log of the partition function:

$$\Gamma_{1\text{PI}}[\phi] = W[J] - \int d^4x J(x)\phi(x), \tag{56}$$

and $W[J] = -i \ln Z[J]$. Often, classical backgrounds of a QFT (meaning solutions to the classical equations of motion) get corrected by loops and can drastically change behaviour. However, the 1PI action takes all of these corrections into account, and solutions to the equations of motion of the effective action tell us the good quantum vacua. Note that the 1PI action is treated as a classical action, meaning it satisfies the classical BV master equation $\{S, S\} = 0$ rather than the quantum master equation (47).

The 1PI action generates, as its name suggests, all 1PI Feynman diagrams. It can be constructed in SFT by simply changing our definition of the string vertices: instead of letting $\mathcal{V}_{g,n}$ be a section over the region of $\mathcal{M}_{g,n}$ not covered by $\{\Sigma_1, \Sigma_2\}$ and $\Delta\Sigma_3$, we will let it be a section over the region not covered by $\{\Sigma_1, \Sigma_2\}$ only. In other words, the loop contributions from $\Delta\Sigma_3$ are being absorbed into the string vertices, just leaving separating Riemann surfaces connected by propagators, i.e. the 1PR graphs. Since the only change is the definition of $\mathcal{V}_{g,n}$, the form of the action is exactly the same, just replacing $\{A_1, \ldots, A_n\}$ with $\{A_1, \ldots, A_n\}_{1\text{PI}}$, which uses the new vertices. Since we will mainly use the 1PI action, we will drop the 1PI subscript. As we will see later in §4.2, the difficulties introduced by $\tilde{\Psi}$ in the superstring field theory mean that it is much easier to prove background independence using this effective action and its equation of motion.

3.5 Backgrounds in string field theory

The 1PI equation of motion for the closed superstring field theory takes the same form as the classical equation of motion, but it uses the 1PI vertex in defining the string bracket $|[A_1, \ldots, A_n]\rangle$:

$$Q_B |\Psi\rangle + \sum_{n=1}^{\infty} \frac{1}{n!} \mathcal{G} |[\Psi^n]\rangle = 0. \tag{57}$$

As mentioned above, solutions to this equation are the backgrounds of the full quantum SFT. The simplest solution is $\Psi = 0$; this is the solution representing the CFT which we used in constructing the SFT.

Hence, there are two different ways of understanding the backgrounds of string theory:

1. A background is given by the worldsheet CFT, which in turn corresponds to some set of background fields in the spacetime, as in (19). Since SFTs are defined using a background CFT, there will be a *distinct* SFT for every possible set of background fields.

2. Within the models of a particular SFT, the different backgrounds are given by solutions to (57).

The relation between these two notions will be critical in interpreting background independence proofs, and we will discuss this relation in more detail in §4 and §5.

However, to get a simple intuitive picture, we can imagine the spacetime as consisting of a manifold M and some background field configuration $\hat{B}$. SFT has fields of all spins, so there are background tensor and spinor fields of all ranks. We will denote such a configuration with the following:

$$\langle M, \hat{B} \rangle := \langle M, \hat{T}, \hat{A}_\mu, \hat{g}_{\mu\nu}, \dots \rangle. \tag{58}$$

The kinematical possibilities of a string field theory are $\langle M, \hat{B}, \Psi \rangle$. The string field Ψ represents fluctuations around $\hat{B}$:

$$\langle M, \hat{B}, \Psi \rangle := \langle M, \hat{T} + [\Psi], \hat{A}_\mu + [\Psi]_\mu, \hat{g}_{\mu\nu} + [\Psi]_{\mu\nu}, \dots \rangle, \tag{59}$$

where $[\Psi]_{\mu_1 \dots \mu_n}$ represent the rank n tensor component of the Fock space expansion of the string field (28).[36]

It is then clear that the solution $\Psi = 0$ of the SFT represents $\hat{B}$, i.e. our initial background, in spacetime. Other dynamical possibilities are given by other solutions of the 1PI equation of motion (57). Consider a specific solution Ψ'. The model representing this solution is

$$\begin{aligned} \langle M, \hat{B}, \Psi' \rangle &= \langle M, \hat{T} + [\Psi'], \hat{A}_\mu + [\Psi']_\mu, \hat{g}_{\mu\nu} + [\Psi']_{\mu\nu}, \dots \rangle \\ &:= \langle M, \hat{T}', \hat{A}'_\mu, \hat{g}'_{\mu\nu}, \dots \rangle. \end{aligned} \tag{60}$$

However, we could have taken this configuration as our *initial* background configuration; let us denote it as $\hat{B}'$. We could have formulated a different string field theory around $\hat{B}'$, say with dynamical variable Π'. Then the kinematical possibilities of this new theory are

$$\langle M, \hat{B}', \Pi' \rangle := \langle M, \hat{T}' + [\Pi'], \hat{A}'_\mu + [\Pi']_\mu, \hat{g}'_{\mu\nu} + [\Pi']_{\mu\nu}, \dots \rangle. \tag{61}$$

The background $\hat{B}'$ is represented by $\Pi' = 0$ in this new string field theory. It therefore seems that the model $\langle M, \hat{B}, \Psi' \rangle$ in the first SFT is the same as the model $\langle M, \hat{B}', 0 \rangle$ in the second; both represent the configuration

$$\langle M, \hat{T}', \hat{A}'_\mu, \hat{g}'_{\mu\nu}, \dots \rangle \tag{62}$$

in spacetime. The background independence proofs we will consider in §4 essentially extend this correspondence to all kinematical possibilities of both theories by constructing an isomorphism between them; moreover, dynamical possibilities will always be taken to dynamical possibilities under such maps. However, these proofs are done at the level of the worldsheet Hilbert space, and the naïve spacetime interpretation presented here likely must be modified, as will be discussed in §5.

[36] This is true at linear order, which is sufficient for the sketch we're presenting here, but in general $[\Psi]_{\mu_1, \dots \mu_n}$ will contain non-linear contributions from other components of the string field, and in fact the precise identification depends on the definition of the string vertices $\mathcal{V}$ discussed above; see Mazel et al. (2025) for more details.

3.6 Witten's cubic theory

The above closed SFTs are extraordinarily complicated, and in particular, only perturbative solutions to the equations of motion are currently known. Fortunately, for open bosonic strings, there is a much simpler theory given by Witten's cubic string field theory (CSFT) (Witten 1986). The relative simplicity of the theory means that certain *non-perturbative* results are known, which will play some role in questioning the naïve spacetime interpretation above.

Open strings end on objects called Dp-branes (henceforth just called D-branes), which live on $p+1$ dimensional submanifolds of the spacetime. The excitations of the open string are then interpreted as fields living on the brane and represent deformations of the brane. From the worldsheet point of view, open strings are Riemann surfaces with boundary. D-branes are then interpreted as different boundary conditions of the worldsheet theory.

To formulate CSFT, we need an open *and* closed string background, meaning we need a spacetime with a set of background fields and a configuration of D-branes on this spacetime with background fields localized on the D-branes. On the worldsheet, this means we need to choose a boundary action in addition to our CFT, which makes it a boundary conformal field theory (BCFT). Note that the boundary action is constructed using operators from the CFT.

Quantum mechanically, open strings can combine into closed strings, meaning a truncation of string theory to only open strings would be inconsistent in the full quantum string theory. For this reason, we will only work at the classical level (in the sense that we only work at tree level), and there will be no closed string fluctuations at all. In particular, the spacetime itself cannot change in this theory (there is no gravity), and it should instead be seen as a theory describing the worldvolume excitations of D-branes. It is more like a stringy generalization of Yang–Mills theory than general relativity.

The cubic theory is typically written using an associative $*$-product and takes the form[37]

$$S = \mathrm{Tr}\left(\Psi * Q_B\Psi + \frac{2}{3}\Psi * \Psi * \Psi\right), \tag{63}$$

which looks like a Chern–Simons action. It can be rewritten in the language of conformal field theory correlation functions,[38] where it takes a form more similar to the closed string field theory:

$$S = \frac{1}{2}\langle\Psi|\, Q_B\,|\Psi\rangle + \frac{1}{3}\langle\Psi,\Psi,\Psi\rangle. \tag{64}$$

This cubic vertex is very special, as it can cover the entire moduli space of all Riemann surfaces (with boundary) by itself! Thus, this theory can reproduce all open string amplitudes, and we do not need any additional vertices. The equations of motion are

$$Q_B\Psi + \Psi * \Psi = 0, \tag{65}$$

[37] For the definition of this associative $*$-product, see Witten (1986).

[38] See Erler (2022, eq. 3.169) for the explicit form of this vertex.

and the gauge symmetry is

$$\Psi \to \Psi + Q_B\Psi + \Psi * \Lambda - \Lambda * \Psi. \tag{66}$$

Solutions to the equations of motion correspond to different backgrounds and are interpreted as different D-brane configurations on the spacetime. The simple solution $\Psi = 0$ is just our starting configuration. As discussed above, we can construct a different CSFT for every background configuration of D-branes.[39]

4 Background independence

With this basic background regarding SFT in hand, we turn now to the claims which are made in the literature regarding its background independence. But we do so obliquely, by first considering a simpler, motivating example, to do with the spin-2 reformulation of general relativity. After considering this case in §4.1, we present a schema for the equivalence of SFTs and their background independence in §4.2, before turning in §4.3 and §4.4 to two different ways of filling out this schema.

4.1 Motivating example: spin-2 gravity

Before we tackle claims regarding the background independence of SFT, it will be instructive to review how such ideas play out in a simpler and more familiar setting: that of a spin-2 reformulation of general relativity.[40] For simplicitly, we'll consider the case of vacuum general relativity. Kinematically possible models of vacuum general relativity are given by tuples $\langle M, g_{\mu\nu}\rangle$, where M is a four-dimensional smooth manifold and $g_{\mu\nu}$ a Lorentzian metric. The dynamically possible models are then picked out by the vacuum Einstein field equations $G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 0$, which can be obtained by varying the Einstein–Hilbert action

$$S_{\text{EH}}[g_{\mu\nu}] = \int_M d^4x\ \mathcal{L}_{\text{EH}}[g_{\mu\nu}] = \int_M d^4x\ \sqrt{-g}R[g_{\mu\nu}], \tag{67}$$

where $R[g_{\mu\nu}]$ is the Ricci scalar associated to $g_{\mu\nu}$ and $g = \det(g_{\mu\nu})$.

Pick now an arbitrary dynamically possible model $\langle M, \hat{g}_{\mu\nu}\rangle$ where $\hat{g}_{\mu\nu}$ is a solution of the Einstein field equations (meaning that $G_{\mu\nu}[\hat{g}_{\rho\sigma}] = 0$) and split the metric $g_{\mu\nu}$ in (67) as $g_{\mu\nu} = \hat{g}_{\mu\nu} + h_{\mu\nu}$, where $h_{\mu\nu}$ is now an arbitrary two-index tensor and is not in general a metric. Remember that $\hat{g}_{\mu\nu}$ is a fixed, particular solution to the Einstein field equations meaning that we treat it as a non-dynamical variable and *don't* (at least initially) subject it to variation via

[39] The presentation here is somewhat anachronistic, since the D-brane interpretation of the solutions of the equations of motion (65) was available only after the discovery of D-branes, with Dai et al. (1989) and Hořava (1989).

[40] The example considered in this section is sketched in Erbin's handwritten notes on BI of SFT (Erbin 2022). In this philosophical literature, spin-2 gravity has also been discussed recently by Linnemann et al. (2023) and Salimkhani (2020); the verions there are not the same as the spin-2 approach considered here.

Hamilton's principle. Consequently, $h_{\mu\nu}$ subsumes all the interesting dynamics and is the new dynamical variable of $S_{\text{EH}}[\hat{g}_{\mu\nu} + h_{\mu\nu}]$ effectively replacing $g_{\mu\nu}$ in this role. Expanding S_{EH} as a power series around $\hat{g}_{\mu\nu}$ in terms of $h_{\mu\nu}$ we obtain

$$\begin{aligned} S_{\text{EH}}[g_{\mu\nu}] &= S_{\text{EH}}[\hat{g}_{\mu\nu} + h_{\mu\nu}] = S_0[\hat{g}_{\mu\nu}, h_{\mu\nu}] + S_1[\hat{g}_{\mu\nu}, h_{\mu\nu}] + \dots \\ &= \int_M d^4x \ \mathcal{L}_{\text{EH}}[\hat{g}_{\mu\nu}] + \int_M d^4x \ \frac{\partial \mathcal{L}_{\text{EH}}[g_{\mu\nu}]}{\partial g_{\mu\nu}}\bigg|_{g_{\mu\nu}=\hat{g}_{\mu\nu}} h^{\mu\nu} + \dots \\ &= \int_M d^4x \ \sqrt{-\hat{g}} R[\hat{g}_{\mu\nu}] + \int_M d^4x \ \sqrt{-\hat{g}} G_{\mu\nu}[\hat{g}_{\mu\nu}] h^{\mu\nu} + \dots \\ &=: S_{\text{spin-2}}[\hat{g}_{\mu\nu}, h_{\mu\nu}]. \end{aligned} \tag{68}$$

Note that if $\hat{g}_{\mu\nu}$ is a solution to the Einstein field equations as we assumed, then $S_1[\hat{g}_{\mu\nu}, h_{\mu\nu}]$ vanishes and we are left only with higher-order terms and S_0.

Imagine now an alternative course of history in which the discoveries of early-twentieth century physics unfolded in quite a different fashion.[41] In this alternative course of events, general relativity wasn't conceived of by Albert Einstein and David Hilbert around 1915, but rather by two very different scientists: Albert Beinstein and David Dilbert. Moreover, the formulation of the theory itself happened to be somewhat different from the Einstein–Hilbert formulation. While the *actual* formulation of general relativity—the one described by the Einstein–Hilbert action S_{EH}—is very elegant in the sense that S_{EH} doesn't require a choice of a fixed background field to be written down, this is not so in the Beinstein–Dilbert formulation (spin-2 theory for short). In this spin-2 theory, the action is given by $S_{\text{spin-2}}$ and as such requires a choice of a fixed, privileged background $\hat{g}_{\mu\nu}$ in order to be written down. In other words, the Beinstein–Dilbert formulation given by $S_{\text{spin-2}}$ appears to be *background dependent* in a loose, to-be-made-precise sense.[42] That said, we note one last interesting property of the relationship between this spin-2 theory and general relativity: the two theories are in an important sense equivalent, because one seems to be just a reformulation of the other since the Einstein–Hilbert action is related to the Beinstein–Dilbert action by a simple change of variables as demonstrated in (68).

With some ingenuity, Beinstein and Dilbert could have noticed that their action is invariant under a special kind of transformation which maps

$$\begin{aligned} &\hat{g}_{\mu\nu} \mapsto \hat{g}'_{\mu\nu}, \\ &h_{\mu\nu} \mapsto h'_{\mu\nu} := h_{\mu\nu} + \hat{g}_{\mu\nu} - \hat{g}'_{\mu\nu}. \end{aligned} \tag{69}$$

Here $\hat{g}'_{\mu\nu}$ is another fixed solution to the Einstein field equations. To Beinstein and Dilbert, this invariance would have presumably seemed like a miraculous

[41] In this passage we of course pay homage to our heroes, Brown (2005) and Stachel (2007).

[42] We set aside the interesting albeit tangential question of why and how Beinstein and Dilbert arrived at the spin-2 theory action given its reliance on a suspect background field, infinitely many terms and overall cumbersome form; after all, questions of this kind are more properly suited for the historians!

conspiracy. Each of the infinitely many terms in $S_{\text{spin-2}}$ has very non-trivial dependence on $\hat{g}_{\mu\nu}$ which hides inside the power series coefficients and each successive term in the power series contains increasingly higher powers of $h_{\mu\nu}$ rendering the symmetry transformations (69) hidden on first blush.[43] However, to we followers of Einstein and Hilbert, this invariance of $S_{\text{spin-2}}$ is revealed to be nothing but a manifestation of a trivial property of the Einstein–Hilbert action. For from (68), it follows immediately that

$$\begin{aligned} S_{\text{spin-2}}[\hat{g}'_{\mu\nu}, h'_{\mu\nu}] &= S_{\text{EH}}[\hat{g}'_{\mu\nu} + h_{\mu\nu} + \hat{g}_{\mu\nu} - \hat{g}'_{\mu\nu}] \\ &= S_{\text{EH}}[\hat{g}_{\mu\nu} + h_{\mu\nu}] = S_{\text{EH}}[g_{\mu\nu}] = S_{\text{spin-2}}[\hat{g}_{\mu\nu}, h_{\mu\nu}], \end{aligned} \tag{70}$$

which is the statement of invariance of $S_{\text{spin-2}}$ under (69). Invariance under (69) is thus hidden deep down in the intricate structure of $S_{\text{spin-2}}$ and but becomes manifest in S_{EH}.[44] Moreover, the fact that such invariance exists allows us to conclude that the choice of $\hat{g}_{\mu\nu}$ in $S_{\text{spin-2}}$ is in some sense irrelevant provided that it is compensated for by appropriate change of the dynamical variables $h_{\mu\nu}$.

We can more formally express this 'change of variables' between two different spin-2 theories as follows. Suppose we have two different theories of spin-2 gravity with two different 'background metrics' $\hat{g}_{\mu\nu}$ and $\hat{g}'_{\mu\nu}$. The dynamical variable in the first theory will be denoted $h_{\mu\nu}$, whereas the dynamical variable in the second theory will be $h'_{\mu\nu}$. Let us define (suppressing indices)

$$\begin{aligned} S_1[h] &:= S_{\text{spin-2}}[\hat{g}, h], \\ S_2[h'] &:= S_{\text{spin-2}}[\hat{g}', h']. \end{aligned} \tag{71}$$

Both S_1 and S_2 are defined on the space of sections of symmetric two-index tensors on the manifold: $S_i : \Gamma(\text{Sym}(T^*M \otimes T^*M)) \to \mathbb{R}$. Now, let a bijection $F : \Gamma(\text{Sym}(T^*M \otimes T^*M)) \to \Gamma(\text{Sym}(T^*M \otimes T^*M))$ (the 'change of variables map') be defined by $F(h) = h - \hat{g} + \hat{g}'$, i.e. (69); then, by the above argument, the pullback of S_2 under this bijection is S_1:

$$F^*S_2[h] := S_2[F(h)] = S_1[h]. \tag{72}$$

Again, this is simply a formal way to say that the two actions are related by a change of variables.

The kinematically possible models of the first theory with action S_1 are $\langle M, \hat{g}_{\mu\nu}, h_{\mu\nu}\rangle$, while those of the latter theory with action S_2 are $\langle M, \hat{g}'_{\mu\nu}, h'_{\mu\nu}\rangle$. These are related to the corresponding model $\langle M, g\rangle$ in general relativity by

$$g_{\mu\nu} = \hat{g}_{\mu\nu} + h_{\mu\nu} = \hat{g}'_{\mu\nu} + h'_{\mu\nu}. \tag{73}$$

We then have the following action-preserving isomorphism of models between the two spin-2 theories:

$$\langle M, \hat{g}, h\rangle \sim \langle M, \hat{g}', h + \hat{g} - \hat{g}'\rangle. \tag{74}$$

[43] For recent philosophical work on hidden symmetries, see Bielińska and Jacobs (2024) and Read (2025).

[44] This is somewhat akin to Trautman gauge symmetry being a 'hidden' symmetry of Newtonian gravitation theory, made manifest in Newton–Cartan theory. See Read (2025).

Since the action is preserved under this map, the dynamically possible models are taken to dynamically possible models in this isomorphism. If this isomorphism really represents a physical equivalence between models of different theories, then presumably only some invariant structure between these models should be taken to be physical. In the case of spin-2 gravity, this is of course given by the metric of general relativity (73) (modulo diffeomorphisms).

We'll soon see that there are many analogies between the situation in SFT and our spin-2 theory toy example. In fact, Sen and Zwiebach's proofs of background independence of SFT—which we take up and analyze in §4.4—can be seen as analogous to the ingenious step of seeing the non-trivial invariance of $S_{\text{spin-2}}$ under (69) *without* the prior knowledge of S_{EH} and its relationship to $S_{\text{spin-2}}$ encapsulated by (68).

4.2 Schema for equivalence of SFTs and BI proofs

Background independence proofs in SFT turn out to have many parallels to the spin-2 case presented above. In particular, the overarching goal is to show that two different SFTs formulated around different backgrounds are actually equivalent. In this section, we will describe a schema for showing when two string field theories are equivalent. This schema will hold for all string field theories, closed or open.

In parallel with the spin-2 case, for two SFTs to be equivalent, we require some kind of isomorphism of models between the two theories,

$$\langle M_1, \hat{B}_1, \Psi_1 \rangle \sim \langle M_2, \hat{B}_2, \Psi_2 \rangle. \tag{75}$$

Here, M_i is the spacetime manifold, $\hat{B}_i$ is the background structure of the SFT, and Ψ_i is the dynamical string field.

We don't just want an arbitrary isomorphism of models, however; this of course can be achieved for any two theories whose solution spaces are of the same dimension/cardinality. Some salient dynamical structure must be preserved under the mapping. In the classical case, by analogy with spin-2 gravity, we require that the actions be equivalent under a pullback, i.e. that the different actions be related by a change of variables. In addition, due to the gauge symmetry of SFT, we also need it to be the case that the BV structures of the two theories match. In the quantum case, we want the map to preserve all correlation functions, the path integral, and the BV structure of the quantum master action. Equivalently, we can preserve the 1PI equations of motion. This will ensure that dynamical possibilities are taken to dynamical possibilities and that the models are empirically equivalent.

One of the key conceptual ideas of the background independence proofs of Sen and Zwiebach (1994a,b) is that, since the two models $\langle M_1, \hat{B}_1, \Psi_1 \rangle$ and $\langle M_2, \hat{B}_2, \Psi_2 \rangle$ will be empirically equivalent (since the actions or correlation functions will be the same), they should *eo ipso* be regarded as being *physically* equivalent.[45] Then, the backgrounds $\hat{B}_1$ and $\hat{B}_2$ have no physical significance;

[45] Readers familiar with the literature on the philosophy of symmetries will recognise this as

presumably, only some invariant structure between these models will be physical, just as with the dynamical metric $g_{\mu\nu}$ in the case of spin-2 gravity and general relativity. We will discuss what this invariant structure could be in §5 below. Note that this is a sort of generalization of the debates about symmetry: there, symmetry-related models *within* some theory are identified, whereas in the SFT case, we wish to identify models *between* two (or more) theories.[46] In this sense, SFT equivalences are something like 'inter-theoretic' gauge equivalences; the different backgrounds and corresponding actions are different 'gauge choices' which are gauge equivalent, albeit not manifestly so.

With all of this in mind, let's now give the general schema for the equivalence of SFTs. Recall that a given SFT, call it SFT_i, is specified by the following data:

1. The worldsheet CFT, which comes equipped with a Hilbert space $\mathcal{H}_i$ and Fock space basis $|\phi_i\rangle$.[47]

2. A dynamical string field $\Psi_i \in \mathcal{H}_i$.

3. The SFT action $S_i : \mathcal{H}_i \to \mathbb{R}$.

Let us now have two string field theories SFT_1 and SFT_2 with an isomorphism (up to gauge transformation[48]) between their Hilbert spaces:

$$f : \mathcal{H}_1 \to \mathcal{H}_2$$
$$f^{-1} : \mathcal{H}_2 \to \mathcal{H}_1.$$

As already mentioned above, the kinematically possible models of a string field theory are determined by (the gauge equivalence classes of) Ψ_i, which is an arbitrary member of $\mathcal{H}_i$. More precisely, the models are given by $\langle M_i, \hat{B}_i, \Psi_i \rangle$, where $\hat{B}_i$ is the background structure determined by the underlying worldsheet CFT; since this is fixed, every distinct model is in one-to-one correspondence with the (gauge equivalence classes of the) members of the Hilbert space. Thus, if we have such an isomorphism of Hilbert spaces, then we already have an isomorphism of models. In order for the theories to be equivalent, however, this isomorphism must preserve the action under a pullback, meaning the two theories are related by a change of variables:

$$f^* S_2[\Psi_1] \equiv S_2[f(\Psi_1)] = S_1[\Psi_1]. \tag{76}$$

something like 'interpretationalism' about symmetries, according to which one can regard empirically equivalent, symmetry-related models of physical theories *ab initio* as being physically equivalent. See Møller-Nielsen (2017) and Read and Møller-Nielsen (2020) for details.

[46] This is the kind of case considered by Read and Møller-Nielsen (2020) in the case of dualities in physics.

[47] We assume $\mathcal{H}_i$ is the part of the more general CFT Hilbert space which satisfies the level matching conditions.

[48] The isomorphism takes gauge equivalence classes to gauge equivalence classes—recall (44). Actually, in general, we only need an isomorphism between cohomologies (gauge equivalence classes) of the theories, and it might not be possible to have a direct isomorphism of Hilbert spaces. This is the case for the Erler–Maccaferri solutions discussed in the next subsection. One can also phrase this requirement as saying that the BV form is preserved, $\omega_1 = f^* \omega_2$.

For two SFTs to be equivalent, we just need there to be an isomorphism up to gauge transformations $f : \mathcal{H}_1 \to \mathcal{H}_2$ that satisfies (76). This will ensure that every model in SFT_1 has an empirically(/physically) equivalent counterpart in SFT_2 and vice versa.

This map f will typically consist of three parts (all isomorphisms up to gauge transformations):

1. A translation $T : \mathcal{H}_1 \to \mathcal{H}_1$.
2. A field redefinition $G : \mathcal{H}_1 \to \mathcal{H}_1$.
3. An abstract Hilbert space isomorphism $F : \mathcal{H}_1 \to \mathcal{H}_2$.

Then, $f = F \circ G \circ T : \mathcal{H}_1 \to \mathcal{H}_2$ will be the action-preserving isomorphism in (76).

The translation T can be thought of as shifting to the solution in SFT_1 that represents the background configuration $\hat{B}_2$ of SFT_2. To find it, one must solve the 1PI equation of motion (57). Let $\hat{\Psi}_2 \in \mathcal{H}_1$ be this solution; then we expand around it (just as in the spin-2 case, where we expand around a particular background metric): $\Psi_1 = \hat{\Psi}_2 + \Phi_1$. This Φ_1 will be our new dynamical variable, and T implements this shift:

$$T(\Psi_1) = \Psi_1 - \hat{\Psi}_2 = \Phi_1. \tag{77}$$

The field redefinition G ensures that the form of the actions is the same for both theories, and the map F will be a general isomorphism between the two Hilbert spaces of the CFTs. In some cases, it can be interpreted as a 'connection on the space of CFTs' (see Sen and Zwiebach (1994b)).

In the quantum case, instead of equating the actions, we equate the whole path integral with measure (Sen and Zwiebach 1994b):

$$\mathcal{D}\Psi_1 e^{S_1[\Psi_1]} = f^*(\mathcal{D}\Psi_2 e^{S_2[\Psi_2]}). \tag{78}$$

This will ensure the equivalence of all correlation functions as long as the gauge equivalence classes are preserved, i.e. if the BV form is preserved. To see this, let A_2 be an operator in $\mathcal{H}_2$ which satisfies the BV condition. Then $A_1 := f^* A_2$ will also satisfy the BV condition in $\mathcal{H}_1$ and will therefore be a physical observable. Moreover,

$$\begin{aligned} \langle A_1 \rangle &= \int_{L_1} \mathcal{D}\Psi_1 A_1(\Psi_1) e^{S_1} \\ &= \int_{f^* L_2} f^* \left(\mathcal{D}\Psi_2 A_2(\Psi_2) e^{S_2} \right) \\ &= \int_{L_2} \mathcal{D}\Psi_2 A_2(\Psi_2) e^{S_2} = \langle A_2 \rangle. \end{aligned} \tag{79}$$

To go from the second to the third line, we note that the pullback simply expresses a change of variables in the path integral. Thus, we find that the

correlation functions of the two theories are equivalent. Note that equating the path integral (78) is identical to equating the 1PI action under a pullback. This can be seen from the definition of the 1PI action (56), which involves performing the full path integral over the classical fields through the dependence on $Z[J]$. Thus, the quantum case works identically to the classical case with the simple replacement of the action with the 1PI action.

Instead of requiring (78), we can follow Sen (2018) in equating the 1PI equations of motion under the following:

$$Q_{B,1}\left|\Psi_1\right\rangle + \sum_{n=1}^{\infty} \frac{1}{n!} \mathcal{G}_1 \left|[\Psi_1^n]\right\rangle = f^{-1}\left(Q_{B,2}\left|f(\Psi_1)\right\rangle + \sum_{n=1}^{\infty} \frac{1}{n!} \mathcal{G}_2 \left|[f(\Psi_1)^n]\right\rangle \right). \tag{80}$$

This will also lead to an empirical equivalence of the two theories, and dynamical possibilities will be taken to dynamical possibilities, since the equations of motion are preserved (which are what classify the dynamical possibilities). Sen's original proof of background independence of closed superstring field theory needed to follow this strategy of equating the equations of motion rather than the actions. The reason is due essentially to the doubled string field $\tilde{\Psi}$; indeed, one can prove that Sen's action (39) *cannot* be background independent in the sense of equivalence of (1PI) actions, see Sen (2018). However, recently, a new action for superstring field theory has been found by Hull (2025). With this new formulation, it is possible to directly equate the 1PI actions, and it is no longer necessary to invoke the equations of motion.

Of course, the above is but a sketch, and in practice the challenge lies in identifying/constructing the isomorphism $f = F \circ G \circ T : \mathcal{H}_1 \to \mathcal{H}_2$. Over the next two subsections, we'll look at implementations of this schema to the open string (§4.3) and closed string (§4.4). In each case, to repeat, what's involved is explicitly constructing this f.

4.3 The Erler–Maccaferri solutions

Let us now implement this schema for the open string. Recall from §3.6 Witten's (classical) cubic action for a CSFT,

$$S_1 = \mathrm{Tr}_1\left(\Psi_1 * Q_{B1}\Psi_1 + \frac{2}{3}\Psi_1 * \Psi_1 * \Psi_1 \right). \tag{81}$$

This action requires background closed *and* open string fields; we will denote them as $\hat{C}_1$ and $\hat{O}_1$ respectively. Consider another CSFT, say with action S_2 and dynamical variable Ψ_2, formulated around a different open string background, but with the same closed string background: $\hat{C}_1 = \hat{C}_2$, $\hat{O}_1 \neq \hat{O}_2$. In other words, consider a CSFT formulated around a different D-brane configuration in the same spacetime.

The condition for background independence (76) requires that there exist a map f between the Hilbert spaces of the two CSFTs which preserves the

action. As we have seen, a typical map f will involve a translation (finding the solution) T, a field redefinition G, and a map F between the Hilbert spaces. For the case of CSFT, the difficult part of proving the existence of f is in finding the solutions corresponding to different backgrounds, i.e. finding T. However, Erler and Maccaferri (2014, 2020) explicitly constructed all *non-perturbative* solutions to the CSFT equations of motion which represent other CSFT backgrounds. There is also a natural map F between the Hilbert spaces of all pairs of CSFTs with the same $\hat{C}$, which is given by open strings connecting the D-branes. The field redefinition G is simply the identity in this case.

If one expands in some CSFT_1 around the Erler–Maccaferri solution $\hat{\Psi}_2^{\mathrm{EM}}$ representing the other CSFT_2 background $\hat{O}_2$, then uses the map F (denoted $F(\Phi) = \bar{\Sigma}\Phi\Sigma$ in this case), one can show that the actions S_1 and S_2 are equivalent under the pullback of $f = F \circ T$: $f^* S_2[\Psi_1] = S_1[\Psi_1]$; moreover, this map preserves gauge equivalence classes under (66). Thus, we have an action-preserving isomorphism between the models of any two CSFTs formulated with the same closed string background:

$$\langle M, \hat{C}; \hat{O}_1, \Psi_1 \rangle \sim \langle M, \hat{C}; \hat{O}_2, \Psi_2 \rangle. \tag{82}$$

If we accept this criterion as amounting to physical equivalence, then it seems that the choice of open string background $\hat{O}_i$ is irrelevant to the physics of CSFT and hence should be given no ontological significance. The theory is still clearly background dependent because of $\hat{C}$, but this argument suggests that some of the other structure needed in formulating the theory is superfluous. In the open string case, it is possible to give a formulation which is manifestly independent of $\hat{O}$, as will be discussed in §5.3.

4.4 The quantum closed string

We will now use the schema for both the bosonic and supersymmetric closed string field theories. These are arguably most interesting from the perspective of background independence, since they are quantum theories where spacetime will also fluctuate. However, due to the complexity of the theories, we lack many of the results present for the open string. In particular, there is no known analogue of the Erler–Maccaferri solutions, as it is simply not known how to solve analytically the closed string field theory equations of motion for backgrounds 'far away' from our starting configuration. However, there is still an isomorphism of models for 'infinitesimally close' backgrounds. From the CFT point of view, these backgrounds are related by small marginal deformations. From the spacetime point of view, the backgrounds $\hat{B}$ and $\hat{B}'$ differ by infinitesimally small field values.

Sen and Zwiebach (1994a,b) and Sen (2018) prove the background independence of bosonic closed string field theory and closed superstring field theory respectively for such 'infinitesimal' shifts in backgrounds. The map T can be found by recursively solving the equations of motion (57) order-by-order in some small parameter ϵ. Then, the map F is given by a *connection on the space of*

CFTs: one can view CFTs related by small marginal deformations as living on a manifold with the coordinates representing coupling constants in front of the various possible marginal deformations. The CFT Hilbert spaces form a bundle over this space, and the connection F is an isomorphism between the Hilbert spaces at different points. The most technically challenging part of the proof is the existence of a G that transforms the 1PI action (or equations of motion) into each other.[49] This is done with an explicit construction that is also a recursive series in orders of ϵ. One must show that there are no topological obstructions to this procedure. Once this is established, the existence of a map f between the Hilbert spaces that preserves the action is shown, and hence there is an action-preserving isomorphism of models

$$\langle M, \hat{B}_1, \Psi_1 \rangle \sim \langle M, \hat{B}_2, \Psi_2 \rangle \tag{83}$$

for infinitesimally close backgrounds $\hat{B}_1$ and $\hat{B}_2$.

Sen and Zwiebach (1994a,b) also give an argument that one can 'integrate' such infinitesimal shifts to get finite shifts, and hence that the above isomorphism should extend to closed SFTs formulated around backgrounds $\hat{B}$ and $\hat{B}'$ that are a finite distance apart. This amounts to solving a certain differential equation in the deformation parameter, and Sen and Zweibach prove that there is no topological obstruction in solving this equation. However, there may potentially be divergences in this process which are interpreted as indicating the breakdown of the variables of SFT_1 in describing the background SFT_2, as discussed by Sen (2018).

5 Invariant structure

The background independence proofs from the previous section suggest that some of the structure used in formulating the various string field theories is superfluous. Explicitly, we find that some models are isomorphic: $\langle M_1, \hat{B}_1, \Psi_1 \rangle \sim \langle M_2, \hat{B}_2, \Psi_2 \rangle$. There is therefore a case to be made—or at least explored—that the physical content of the string field theory is whatever is invariant under these isomorphisms. In this section, we will explore various candidates for what amounts to such 'background independent' structure of SFT.

5.1 Invariant (super)gravity structure

Let us focus on the closed string for now. As discussed in §4.4, we have an isomorphism of models between SFTs formulated around different backgrounds $\hat{B}_1$ and $\hat{B}_2$,

$$\langle M, \hat{B}_1, \Psi_1 \rangle \sim \langle M, \hat{B}_2, \Psi_2 \rangle. \tag{84}$$

To understand what this means in more detail, let us consider the spacetime interpretation of a general state in the Hilbert space of one of the string field

[49] As mentioned above, Sen originally needed to equate equations of motion rather than actions due to the auxilliary field $\tilde{\Psi}$. Due to Hull (2025) this is no longer necessary, and one can simply equate the 1PI actions.

theories, $\Psi_1 \in \mathcal{H}_1$. As discussed in §3.5, this will correspond to some set of fields in spacetime, which we can view as a perturbation around the initial background $\hat{B}_1$; in the notation of §3.5, the spacetime model is

$$\langle M, \hat{B}_1, \Psi_1 \rangle = \langle M, \hat{B}_1 + \Psi_1 \rangle := \langle M, \hat{T}_1(x) + [\hat{\Psi}_1(x)], \hat{g}_{1\mu\nu}(x) + [\hat{\Psi}_1(x)]_{\mu\nu}, \dots \rangle. \tag{85}$$

The isomorphism of models between the string field theories (84) tells us that for every spacetime configuration of fields in SFT_1, there is a corresponding empirically equivalent model in SFT_2. This empirically equivalent model will have a similar spacetime representation as

$$\langle M, \hat{B}_2 + \Psi_2 \rangle = \langle M, \hat{T}_2(x) + [\Psi_2(x)], \hat{g}_{2\mu\nu}(x) + [\Psi_2(x)]_{\mu\nu}, \dots \rangle. \tag{86}$$

For these models to be equivalent, it is natural to equate the fields of each rank up to gauge transformations (which we denote by '$\simeq$'):

$$\hat{T}_1(x) + [\hat{\Psi}_1(x)] \stackrel{?}{\simeq} \hat{T}_2(x) + [\hat{\Psi}_2(x)] \tag{87}$$

$$\hat{g}_{1\mu\nu}(x) + [\hat{\Psi}_1(x)]_{\mu\nu} \stackrel{?}{\simeq} \hat{g}_{2\mu\nu}(x) + [\Psi_2(x)]_{\mu\nu} \tag{88}$$

$$\vdots$$

The latter equation (88) bears a striking resemblance to (73) for spin-2 gravity. This similarity suggests that the background independent structure of a model of SFT will be a combination of the background field with the fluctuation, which will simply be the value of the spacetime field $g_{\mu\nu}(x) := \hat{g}_{1\mu\nu}(x) + [\Psi_1(x)]_{\mu\nu} = \hat{g}_{2\mu\nu}(x) + [\Psi_2(x)]_{\mu\nu}$. This will also be the case for the other fields with different ranks, i.e., the background independent model will consist of the sum of the background with the appropriate component of the string field, which will give us the value of the spacetime field. Thus, the background independence proofs suggest that the models of string field theory should be taken to be of the form

$$\langle M, \hat{B} + \Psi \rangle = \langle M, \hat{T} + [\Psi], \hat{g}_{\mu\nu} + [\Psi]_{\mu\nu}, \dots \rangle = \langle M, T(x), g_{\mu\nu}(x), \dots \rangle. \tag{89}$$

However, note that the latter models are precisely the same as those of a gravitational theory with a matter sector which comes from the various string modes. In the case of the bosonic string, these will be a 26D dilaton gravity theory with massive fields of higher spins (plus the tachyon), and in the case of the superstring, these will be the models of the appropriate supergravity extended by similar stringy massive fields.

It therefore seems that the invariant content of the models of SFT is really that of (super)gravity with the matter content being given by the various other string excitations. That said, these models are slightly different from classical, minimally coupled (super)gravity due to quantum corrections coming from stringy interactions. In this sense, SFT is 'just' an extended, UV complete version of spin-2 gravity.

There are, however, some reasons to call into question this natural, invariant (super)gravity structure. In particular, the background independence proofs

relied heavily on the Hilbert spaces of the worldsheet CFTs. Although these Hilbert spaces have Fock space bases which allow straightforward spacetime interpretations of models *within* a specific theory, they can lead to some puzzling results for spacetime interpretations when comparing Fock space bases *between* theories, particularly when comparing backgrounds that are 'far away'. However, this is precisely what we are interested in when assessing background independent structure. To illustrate this point, we will construct the analogous 'natural structure' for Witten's cubic string field theory and discuss some issues that occur when considering the Erler–Maccaferri solutions.

5.2 Invariant Yang–Mills structure

Let us try to apply analogous reasoning from above to the open string field theory. Open string fields living on D-branes are localized to the submanifold of the D-brane. Thus, for example, if we have a D24-brane on a submanifold $\Xi \subset \mathbb{R}^{25,1}$ of 26D Minkowski space, the fields of the brane will be localized to Ξ. The CSFT formulated on this brane describes open string fluctuations of the brane. We will use x as a coordinate on all of Minkowski space and ξ as a coordinate on Ξ. The background of our CSFT will be some collection of closed and open fields; we will separate them with a semicolon for clarity, with closed background fields on the left and open background fields on the right:

$$\langle M, \hat{C}; \hat{O} \rangle = \langle M, \hat{T}(x), \hat{g}_{\mu\nu}(x), \dots; \hat{T}(\xi), \hat{A}_\mu(\xi), \dots \rangle. \tag{90}$$

The open string field Ψ will leave the closed string background unchanged, but will describe fluctuations of the fields living on the brane. Thus, a general model of this CSFT will be

$$\langle M, \hat{C}; \hat{O}, \Psi \rangle = \langle M, \hat{T}(x), \hat{g}_{\mu\nu}(x), \dots; \hat{T}(\xi) + [\Psi(\xi)], \hat{A}_\mu(\xi) + [\Psi(\xi)]_\mu, \dots \rangle. \tag{91}$$

Using the same reasoning as above, the natural background independent structure, or rather, the structure independent of the open string background, should be the sum of the open string background and the fluctuation:

$$\begin{aligned} \langle M, \hat{C}; \hat{O} + \Psi \rangle &:= \langle M, \hat{T}(X), \hat{g}_{\mu\nu}(X), \dots; \hat{T}(\xi) + [\Psi], \hat{A}_\mu(\xi) + [\Psi]_\mu, \dots \rangle \\ &:= \langle M, \hat{T}(X), \hat{g}_{\mu\nu}(X), \dots; T(\xi), A(\xi), \dots \rangle. \end{aligned} \tag{92}$$

In other words, the 'open string background independent structure' should be the values of the fields living on the brane, $T(\xi), A_\mu(\xi), \dots$. This is also the field content of a 24D Yang–Mills theory, again with the appropriate matter sector coming from higher string modes.

The problem arises when we compare this model to a different CSFT which is formulated around some other D-brane configuration with the same closed string background. Let us consider the analogous 'background independent structure' for a CSFT formulated around some D23-brane which lives on a submanifold $\Sigma \subset \Xi$. The fields of the D23-brane will be localized to Σ, which has coordinates σ. Thus, the models are

$$\langle M, \hat{C}; \hat{O}' + \Psi' \rangle = \langle M, \hat{T}(x), \hat{g}_{\mu\nu}(x), \dots; T'(\sigma), A'_\mu(\sigma), \dots \rangle. \tag{93}$$

These are the models of a 23D Yang–Mills theory with massive stringy fields.

This is perhaps not surprising: the effective theory on the worldvolume of a Dp-brane is simply p-dimensional Yang–Mills theory. CSFT is, in a sense, a 'UV completion' of this Yang–Mills theory. However, the Erler–Maccaferri solutions make the rather astonishing suggestion that all CSFTs with the same closed string background are equivalent, and hence, all of these various Yang–Mills theories can arise as effective theories from a CSFT of naïvely different dimension, and even different gauge group. For instance, the simplest configuration of the D23-brane with no excitations (representing a vacuum of the 23D Yang–Mills theory/CSFT) arises as a 'lump' solution within the 24D Yang–Mills theory/CSFT. Moreover, multi D-brane solutions (which can represent $U(N)$ gauge theories) can arise as solutions of the CSFT of a single brane, which only has $U(1)$ gauge symmetry. Thus, the Erler–Maccaferri solutions seem to suggest that even the spacetime fields are not the correct background independent variables for string field theory.

In the closed string case, analogous solutions that are 'far away' are not known, so the spacetime field approach seems more plausible as an account of background independent structure. However, the above argument from the open string suggests that the supergravity models are not sufficient for the full closed SFT, especially when identifying models which may have different matter fields or spacetime topologies. It is therefore important to look at other candidates for background independent structure.

5.3 A theory on the space of theories: Witten's BSFT

In this section, we will look at a formulation of open string field theory which seems to be manifestly independent of the superfluous background structure from earlier; this formulation is called the *boundary string field theory* (BSFT) (Witten 1992).

String fields take values in the Hilbert space of the worldsheet theory $\mathcal{H}$. Moreover, the background independence proofs from above also rely on isomorphisms between these worldsheet Hilbert spaces. Thus, it is plausible that the background independent structure of SFT could be formulated in terms of worldsheet data rather than spacetime data. Indeed, there is a formulation of open string field theory, which in a precise sense is a 'theory on the space of worldsheet theories'. Unfortunately, it is not clear how well-defined this space is due to worries about UV divergences and non-renormalizable boundary interactions. Nonetheless, as we will see, it is possible to write down an action on this space which is manifestly independent of the open string background $\hat{O}$, though it still depends on a fixed closed string background.

Recall that open string field theory is a classical theory, which means we ignore string loops. The worldsheet of an open string without loops is conformally equivalent to a disk D. Consider a worldsheet theory with an action

$$S = \int_D d^2\sigma \, (\hat{g}^{\mu\nu}\partial_\mu X^a \partial_\nu X_a + \dots) + \int_{\partial D} d\sigma \mathcal{V}. \tag{94}$$

This is specified by some closed string background $\langle \hat{g}_{\mu\nu}, \dots \rangle$ and a boundary Lagrangian $\mathcal{V}$, which is a local, ghost number zero operator constructed from the 'bulk' CFT.[50] As with the CSFT, we will work with a fixed closed string background, meaning we will fix the first term of the above action.

Note that in general, the boundary term $\mathcal{V}$ can break conformal invariance. Imposing conformal invariance of the boundary action gives us classical equations for open string background fields, just like how imposing conformal invariance of the 'bulk' CFT gives us equations for closed string background fields, as discussed in §2.2.3. In particular, instead of the Einstein field equations, we get Maxwell's equations (Abouelsaood et al. 1987).

Now, notice that the space of possible boundary Lagrangians $\mathcal{V}$ is just the space of ghost number zero operators of the bulk CFT. Thus, in this sense, the ghost number zero operators of the bulk CFT parameterize the space of possible boundary theories for a given closed string background. To see this more explicitly, one can find a basis of this space V_i and expand the general Lagrangian as $\mathcal{V} = \sum_i \lambda_i V_i$; the coefficients λ_i are coupling constants, and they are the variables of the BSFT. The goal would be to write down an action on this space so that the 'on-shell' $\mathcal{V}$ are the classical open string backgrounds. For technical reasons, however, it is not possible to write an action directly on the space of ghost number zero operators. One must move to the space of ghost number 1 operators $\mathcal{O}$ such that $\mathcal{V} = b_{-1}\mathcal{O}$, where b_{-1} is defined in terms of the bulk b ghost. This will still give us all possible (matter) boundary actions if we choose $\mathcal{O} = c\mathcal{V}$, but there is some redundancy, as multiple $\mathcal{O}$ can represent the same $\mathcal{V}$.[51]

On the space of ghost number 1 operators, Witten's action is implicitly defined by

$$dS_{\text{BSFT}} = \frac{1}{2} \oint \oint d\sigma_1 d\sigma_2 \langle d\mathcal{O}(\sigma_1) \{Q_B, \mathcal{O}(\sigma_2)\} \rangle, \tag{95}$$

where Q_B is the bulk BRST charge, and d is an exterior derivative on the space of couplings in the boundary Lagrangian. The extremal configurations ($dS = 0$) are those that satisfy $\{Q_B, \mathcal{O}\} = 0$. These in turn are precisely the conformally invariant boundary actions (at least in cases where ghosts and matter decouple). Thus, Witten's action tells us that the classical backgrounds of open string theory are given by conformally invariant boundary theories! It is exactly this requirement of worldsheet conformal invariance which gives us the spacetime equations of motion. Also, notice that this action manifestly does not refer to any open string background $\hat{O}$, unlike the cubic string field theory from earlier.

Ultimately then, there's a case to be made that what we have here is a *manifestly* background independent formulation of open SFT; the analogue of

[50] By the bulk CFT, we mean the 2D CFT on the disk specified by the first term in the worldsheet action, i.e., the closed string background. We are not talking about holographic bulk/boundary theories.

[51] For example, this can be done by shifting $\mathcal{O} \to \mathcal{O} + \Delta$, where $b_{-1}\Delta = 0$.

the Einstein–Hilbert action for general relativity in the spin-2 gravity case.[52] This, of course, goes beyond the invariant structures offered in the previous two subsections.

5.3.1 Models of BSFT

The kinematic possibilities of the BSFT are spanned by the ghost number 1 operators of the bulk CFT. However, we will represent the models in terms of the boundary Lagrangian $\mathcal{V}$ directly, since the interpretation of the ghost number 1 operators $\mathcal{O}$ remains somewhat obscure, and we can always go from the ghost number operator to a boundary Lagrangian using $\mathcal{V} = b_{-1}\mathcal{O}$.

In the case where the bulk CFT is a non-linear sigma model ('NLSM'), $\mathcal{V}$ will represent some D-brane configuration within the spacetime, perhaps with some non-zero value for worldvolume fields within the brane. We will simply denote the models in this case as $\langle M, \hat{C}, \mathcal{V}\rangle$, where the operator $\mathcal{V}$ is taken to represent some D-brane configuration within M.

5.4 The cZ action

BSFT liberates open string field theory from the requirement of an initial D-brane configuration by living on the space of worldsheet theories; the equation of motion of the BSFT action tells us that on-shell actions are conformally invariant. It is natural to ask whether the same move can be made for the closed string. That is, can we write down an action on the space of closed string worldsheet theories such that the equations of motion demand conformal invariance?

This problem is technically more challenging than the open string case. In the latter, the bulk closed string CFT provided an 'anchor' so that the space of boundary Lagrangians had a somewhat concrete description in terms of operators in the bulk CFT. On the other hand, the space of possible closed worldsheet theories is the space of all 2D quantum field theories! There are similar worries about the well-definedness of this space due to UV divergences and non-renormalizable field theories.

Despite these difficulties, precisely such an action has been found by Ahmadain et al. (2024) for the classical bosonic closed string.[53] It has a rather unusual form: it is the product of the Zamolodchikov C-function (on a disk of radius r^*) and the partition function of the QFT on the sphere,[54]

$$S_{cZ} = c_{\text{pl}}(r^*) Z_{S^2}. \tag{96}$$

The partition function enforces the conformal invariance condition, and the C-function ensures that the total central charge of the worldsheet theory is $c = 0$

[52] The explicit relation between BSFT and CSFT is explained in Chiaffrino and Sachs (2019) and Totsuka-Yoshinaka (2025).

[53] It is classical for similar reasons to the BSFT: It only contains the sphere partition function of the string, and so it does not account for string loops.

[54] The Zamolodchikov C-function is a function which is minimised on CFTs and which returns there the central charge. For further details, see Ahmadain et al. (2024).

so that there is no conformal anomaly.[55] Both of these functions are defined for all QFTs.

5.4.1 Models of the cZ action

The space of 2D QFTs is difficult to characterize, and the interpretation of a string with an arbitrary QFT as a worldsheet theory is even more poorly understood. For this reason, we will simply restrict our attention to NLSM worldsheet theories such as those in (19). Just like the BSFT above, the variables of the cZ action are the worldsheet couplings themselves. Thus, the models of the cZ action, when restricted to this class of QFTs, can be characterized by the values of the spacetime fields which appear as couplings, $\langle M, T, g_{\mu\nu}, B_{\mu\nu}, \Phi, \dots \rangle$. These are exactly the same models that were inferred in §5.1 from the background independence proofs; as mentioned there, they are the models of 26D dilaton gravity with a stringy matter sector.

5.5 Models of string field theory

As we have seen in this section, there are many ways in which one can characterize the models of SFT; these we have summarized in table 1. All of these representations have benefits and drawbacks.

The naïvely background dependent models are what are most straightforwardly read from the actions of the theories. However, as we know from formulations of gauge theories in which the symmetries are not apparent, it is not always the case that the obvious models are the physically appropriate ones. In particular, the background independence proofs from §4 seem to suggest that certain combinations of the background fields and the string field are the true physical variables; we have called this the 'inferred BI' structure. In some cases, further evidence for these models is provided by the theories in the next column, which give the same dynamics in a way that uses only these BI variables. Notably, however, such a theory is not yet available for the quantum closed string, which is ultimately the most interesting, as it is the only theory of quantum gravity in this table.

6 Assessing the background independence of string field theory

Having now presented and unpacked the relevant details regarding both SFT and the core 'proofs' of the background independence of SFTs, we turn now to a conceptual appraisal of the background independence of SFTs. To do so, we first review some key definitions of background independence from the philosophical literature (drawing upon a more exhaustive survey by Read (2023)) (§6.1); next, we apply these definitions to our toy case study of spin-2 gravity (§6.2); finally,

[55] It is assumed that the ghost sector is not deformed, so the matter CFT must have $c = 26$.

Theory	**Naïvely BD form**	**Inferred BI form**	**Manifestly BI theory**	**Manifestly BI form**
Spin-2 gravity	$\langle M, \hat{g}_{\mu\nu}, h_{\mu\nu} \rangle$	$\langle M, \hat{g}_{\mu\nu} + h_{\mu\nu} \rangle$	EH	$\langle M, g_{\mu\nu} \rangle$
Open SFT	$\langle M, \hat{C}, \hat{O}, \Psi \rangle$	$\langle M, \hat{C}, \hat{O} + \Psi \rangle$	BSFT	$\langle M, \hat{C}, \mathcal{V} \rangle$
Classical bosonic closed SFT	$\langle M, \hat{C}, \Psi \rangle$	$\langle M, \hat{C} + \Psi \rangle$	cZ (NLSM only)	$\langle M, T, g_{\mu\nu}, B_{\mu\nu}, \Phi, \dots \rangle$
Quantum closed SFT	$\langle M, \hat{C}, \Psi \rangle$	$\langle M, \hat{C} + \Psi \rangle$	???	???

Table 1: For each of (i) the spin-2 theory (our running toy example), (ii) the open SFT, (iii) the classical bosonic closed SFT, and (iv) the quantum closed SFT, (a) the naïve background dependent ('BD') structure; (b) the inferred background independent ('BI') structure; (c) the manifestly BI theory; and (d) the manifestly BI structure. By 'manifestly', we here mean that dynamics (typically, an action) for the theory is provided which makes no reference to background variables or fields.

we apply the definitions to the full case of SFT, drawing upon all of the results which we have presented above (§6.3).

6.1 Accounts of background independence

In assessing the background independence of SFT, we draw from the extensive literature which deals with the desideratum of background independence itself. One obvious and urgent question is this: *what is* background independence? The criterion certainly captures a kind of shared intuition in the physics community, but attempts at making this intuition precise have been met with difficulties. A thorough investigation of possible answers to the question of what background independence amounts to has recently been carried out by Read (2023), and in the following we'll draw heavily on the survey of the different accounts of background independence offered in that work.

Focusing initially on the realm of classical theories, Read (2023, ch. 3) identifies a large number of different proposals for analysing the notion of background independence which have been floated in the literature. In the interests of constraining the narrative in this article while still being able to make some interesting points, we'll focus here on the following approaches to background independence:

1. Absolute objects (AO)
2. Variational principles (VP)
3. Belot's proposal (BP)

We now briefly introduce each of these in turn; however, we make no claim to cover these accounts in any degree of completeness. The reader interested in details of the presented accounts (such as ramifications and potential problem cases) should consult Read (2023, ch. 3).

6.1.1 Absolute objects

Read (2023, ch. 3), following Anderson (1967) and later Friedman (1983), defines an absolute object (AO) as "[a] geometric object which is the same (up to isomorphism) in all [dynamically possible models] of a theory." (It's important to be clear here that Read has in mind the 'local' version of an AO proposed by Friedman (1983), rather than the earlier 'global' version proposed by Anderson (1967).) The AO criterion of background independence is then captured by the following definition (Read 2023, p. 15):

Definition 1 (**Background independence, absolute objects**)**.** A theory is background independent iff it has no absolute objects in its formulation.

One of the main problems facing the AO account is that general relativity itself turns out to have an absolute object according to Anderson's definition—the metric determinant—and consequently does not qualify as background independent according to the above definition (see Read (2023, pp. 16–19) for more on this point; cf. Pitts (2006)).

Given these problems with the AO proposal, there are a couple of modifications which are worth mentioning. The first has to do with the notion of a 'fixed field' (terminology introduced by Pooley (2017)). According to Read (2023, p. 12), "for a fixed field [...] the field values at each manifold point are identically the same in every [kinematically possible model]." This is in contrast with absolute objects introduced above which are identical only up to isomorphism and only across dynamically possible models. The fixed field criterion of background independence is then captured by the following definition (Read 2023, p. 20):

Definition 2 (**Background independence, fixed fields**)**.** A theory is background independent iff it has no formulation which features fixed fields.

Note the crucial reliance of the fixed field criterion on the notion of theory formulation. In order for the criterion to be applicable to concrete cases, one needs to supplement it with an understanding of what it means for two theories to be equivalent. Read (2023, ch. 3) considers such criteria; however, for now we only note that some such criteria are required in order to reach definite conclusions using the fixed field definition.

A third related notion is that of an 'absolute field'. Read (2023, p. 22) defines an absolute field as "[a] geometric object specified in the [kinematically possible models] of a theory, which is fixed (up to isomorphism) in all [dynamically possible models] of that theory." This is in contrast with absolute objects (e.g. the metric determinant) which *don't* have to be specified explicitly in the models of the theory. The absolute field criterion of background independence is then captured by the following definition (Read 2023, p. 22):

Definition 3 (**Background independence, absolute fields**)**.** A theory is background independent iff it has no absolute fields.

As discussed by Read (2023, ch. 3), the above qualification regarding the metric determinant in general relativity allows the absolute fields criterion to escape the problem case facing the absolute objects criterion; as such, it fares better than the latter in this respect, albeit at the cost of privileging specific formulations of a theory in a way which might seem *ad hoc*.

6.1.2 Variational Principles

The account of background independence in terms of variational principles due to Pooley (2017) differs from the ones covered so far in that it places central emphasis on the action of the theory under consideration. The proposal can be stated as follows (Read 2023, p. 24):

Definition 4 (**Background independence, variational**)**.** A theory is background independent iff its solution space is determined by a generally covariant action, (i) all of whose dependent variables are subject to Hamilton's principle, and (ii) all of whose dependent variables represent physical fields.

Clause (ii) here plays a crucial role since, in principle, *any* equation of motion can be introduced to a theory's action using Lagrange multipliers. If this proposal is to fare any better than the accounts previously canvassed, it at least needs to be able to tell apart theories which contain genuine background fields from the ones which 'hide' these background fields by defining them at the level of kinematical possibilities, but subjecting them to artificial constraint equations which universally kill their variation across dynamical possibilities. Such artificial constraint equations would usually enter the action principle via unphysical Lagrange multipliers and for this reason (ii) is required. (For further discussion, see Read (2023, ch. 3).)

6.1.3 Belot's proposal

Belot's proposal belongs among the more involved accounts of background independence and it comes with several conceptual novelties compared to the approaches considered so far. In particular, Belot (2011) notes that on his account:

> 1. Background-dependence and independence come in degrees: some theories are fully background-(in)dependent, others nearly so—and others fall somewhere in between.
> 2. A theory can fail to be fully background-independent in virtue of asymptotic boundary conditions.
> 3. The extent of the background-(in)dependence of a theory is not a strictly formal one: in particular, it depends on how one thinks of the geometric structure of each solution and on what sorts of differences between solutions one takes to be unphysical.
>
> (Belot 2011, pp. 2872–3)

The account itself proceeds by first drawing the distinction between *geometrical* and *physical* degrees of freedom. In alignment with the third moral on the above list, Belot avers that whether a particular piece of structure in a physical theory qualifies as geometrical or physical is not strictly a formal matter and our interpretive convictions factor strongly into such judgements. In principle, one can take this to be an advantage of Belot's proposal; however, one should also bear in mind that definitive judgements about the background independence of a particular theory can only be drawn in conjunction with these additional interpretive judgements. Without further ado, let us now introduce the distinction between geometrical and physical degrees of freedom, following the exposition by Read (2023, pp. 28–30).

Consider a theory $\mathcal{T}$ whose kinematically possible models are tuples of the form $\langle M, O^G, O_1, \ldots, O_n \rangle$, where M is a smooth manifold, and O^G is a piece of structure identified antecedently as being 'geometrical'. Such O^G may in principle vary across kinematically and dynamically possible models of $\mathcal{T}$, so let us call $\mathcal{G}$ the set of all O^G across all dynamically possible models of $\mathcal{T}$. Additionally, one may wish to regard certain elements of $\mathcal{G}$ as *equivalent* geometrical structures and equip $\mathcal{G}$ with an equivalence relation $\sim_G$ relating equivalent geometrical structures. Taking the quotient $\mathcal{G}/\sim_G$ then defines the reduced set of geometrical objects in $\mathcal{T}$ called $\tilde{\mathcal{G}}$. Read then notes that "[t]he degrees of freedom needed to parametrize this latter set $\tilde{\mathcal{G}}$ are what Belot calls the *geometrical degrees of freedom* of $\mathcal{T}$." (Read 2023, p. 29)

Turning now to physical degrees of freedom, we note that it is often the case that a given physical theory commits to excess mathematical structure in the sense that two mathematically distinct models may be regarded as having exactly the same representational capacities.[56] Such models are called *gauge-equivalent* and physical theories which contain such models are called *gauge theories*.[57] We denote the relation of gauge equivalence by $\sim_P$, where P in the subscript stands for *physical* equivalence. Belot's procedure for identifying physical degrees of freedom runs as follows. Take the set $\mathcal{D}$ consisting of all the dynamically possible models of $\mathcal{T}$ and take the quotient $\tilde{\mathcal{D}} = \mathcal{D}/\sim_P$. Then "the degrees of freedom needed to parametrize $\tilde{\mathcal{D}}$—the gauge-quotiented class of dynamically possible models is $\mathcal{T}$—are the theory's *physical degrees of freedom*." (Read 2023, p. 29).

Having armed ourselves with the distinction between geometrical and physical degrees of freedom, we can now formulate Belot's four definitions which characterise the background independence of a given theory:

Definition 5 (**Full background dependence, Belot**)**.** A field theory is fully background dependent if it has no geometrical degrees of freedom: every solution is assigned the same spacetime geometry as every other solution.

[56] There are subtleties here regarding the relations between mathematical equivalence and representational equivalence—see Fletcher (2020). We elide them.

[57] Of course, this is just one sense in which the term 'gauge theory' might be used. See Weatherall (2016) for disambiguation.

Definition 6 (**Full background independence, Belot**)**.** A field theory is fully background independent if all of its physical degrees of freedom correspond to geometrical degrees of freedom: two solutions correspond to the same physical geometry iff they are gauge equivalent.

Definition 7 (**Near background dependence, Belot**)**.** A field theory is nearly background dependent if it has only finitely many geometrical degrees of freedom: quotienting the space of geometries that arise in solutions of the theory by the relation of geometrical equivalence yields a finite-dimensional space.

Definition 8 (**Near background independence, Belot**)**.** A field theory is nearly background independent if it has a finite number of non-geometrical degrees of freedom: there is some N such that for any geometry arising in a solution of the theory, the space of gauge equivalence classes of solutions with that geometry is no more than N-dimensional.

Again, this presentation will suffice for our purposes; see Read (2023, ch. 3) for a detailed assessment and exploration of Belot's proposal. We'll see in the remainder of this section that Belot's proposal can be put to very interesting interpretive work when we asses the background independence both of the spin-2 theory and of SFT.

6.2 Background independence of spin-2 gravity

With the above definitions in hand, we'll consider now the background independence of spin-2 gravity, as a warm-up to assessing in the next subsection the background independence of SFT itself.

6.2.1 Models of the spin-2 theory

Having now reviewed some mainstream proposals for what background independence might consist in (as already discussed, this list is non-exhaustive; see Read (2023, ch. 3) for further proposals), we can lay these proposals against our first case of interest: the spin-2 theory. Following up on §4.1, we define the spin-2 theory as follows. Its kinematical possibilities consist of tuples $\langle M, \hat{g}_{\mu\nu}, h_{\mu\nu}\rangle$, where M is a smooth manifold, $\hat{g}_{\mu\nu}$ is a Lorentzian metric and a fixed solution to the Einstein field equations on M, and $h_{\mu\nu}$ is a rank-2 tensor field on M. The dynamically possible models are then picked out by the dynamical equations following from Hamiltonian variation with respect to $h_{\mu\nu}$ of the action functional $S_{\text{spin-2}}$ which we defined in (68). Explicitly, from this variational procedure we obtain the equations of motion

$$0 = \frac{\delta S_{\text{spin-2}}}{\delta h_{\mu\nu}} \implies G_{\mu\nu}[\hat{g}_{\mu\nu}] + O(h_{\mu\nu}) = 0. \tag{97}$$

Since $\hat{g}_{\mu\nu}$ is chosen such that it solves the Einstein field equations on M, we see that the zeroth order term in (97) vanishes. The remaining terms are then proportional to ever-increasing powers of $h_{\mu\nu}$, meaning that $h_{\mu\nu} = 0$ will always

be a solution of spin-2 theory, albeit a trivial one. Formulated in this way, the spin-2 theory by construction includes the entire solution space of general relativity. Note that, on this way of thinking (which we'll adopt in the remainder of this subsection), each choice of background $\hat{g}_{\mu\nu}$ yields a *distinct* spin-2 theory.

6.2.2 Background independence of the spin-2 theory

What should we make of the background independence of the spin-2 theory? On definition 1 (in terms of absolute objects), the theory comes out background *dependent*, since it has an absolute object: the metric determinant $\det(\hat{g}_{\mu\nu})$ is identical across all dynamically possible models of the theory, which means that it is an absolute object in Friedman's sense. (As Read (2023) notes, this is a ubiquitous and problematic feature of the definition in terms of absolute objects: the metric determinant is an absolute object also in general relativity itself, as well as in many other spacetime theories.) Indeed, on our above way of understanding the spin-2 theory (on which each spin-2 theory has its own background $\hat{g}_{\mu\nu}$), the fixed metric $\hat{g}_{\mu\nu}$ *in toto* is also an absolute object (as well as an absolute field, since $\hat{g}_{\mu\nu}$ is specified explicitly in the models of the theory—meaning that definition 3 is also violated).[58,59]

Let us therefore see how the remaining definitions deal with this theory. Turning to definition 2, we note that whether or not $\hat{g}_{\mu\nu}$ is a fixed field will depend upon whether or not it is taken to be fixed *identically* in all kinematical possibilities. If it is not, then this definition does not vindicate any intuitions which one might have regarding the fixed nature of $\hat{g}_{\mu\nu}$.

Definition 4 in terms of variational principles captures a clear sense in which $\hat{g}_{\mu\nu}$ is a fixed background. Recall that the spin-2 action is given by (68) and in deriving the equations of motion (97) only $h_{\mu\nu}$ is subject to Hamiltonian variation. Therefore, spin-2 theory comes out as background dependent on definition 4.

Belot's proposal captured in definitions 5–8 deserves some unpacking in the context of the spin-2 theory. Recall that in order to put the proposal to work, one needs to identify the geometrical and physical degrees of freedom of the theory. Even at this stage, there are decisions to be made when it comes to assessing the background independence of the spin-2 theory with respect to Belot's definitions. For are the fields $\hat{g}_{\mu\nu}$ to be counted as geometrical, and are

[58] If we were to have a more expansive view of the spin-2 theory according to which different, non-isomorphic backgrounds $\hat{g}_{\mu\nu}$ were permitted, then this object would not count as an absolute object—but $\det(\hat{g}_{\mu\nu})$ would still so count. Arguably, this would not be a satisfactory result, for there is a clear sense in which $\hat{g}_{\mu\nu}$ is a fixed background—namely, in the sense that it does not vary within any model, and is not coupled to any other fields. This type of case is discussed by Read (2023, ch. 3) in the context of e.g. a cosmological constant which can vary from model to model, but which is fixed within any particular model.

[59] Are things like the Euler characteristic absolute objects? This would give the spin-2 theory an absolute object. However, this deepens the tension between including other topologies and trivializing background dependence: if there are only models of one topology, we will always have topological invariants. If not, then the fields live in completely different spaces, and it seems impossible for models of manifolds with different topologies to have an isomorphic object, thus eliminating all possibility for an absolute object.

the spin-2 fields $h_{\mu\nu}$ also to be so regarded? While it seems fairly straightforward to regard $\hat{g}_{\mu\nu}$ as being 'geometrical', there's surely a case to be made on either side when it comes to the spin-2 field $h_{\mu\nu}$, for on the one hand (with the origins of the spin-2 theory in general relativity in mind) the object is geometrical, but on other it (of course) looks akin to a spin-2 material field.[60] And even if one does regard $h_{\mu\nu}$ as being geometrical, there are questions to be asked about whether one should be assessing the background independence of the spin-2 theory with $\hat{g}_{\mu\nu}$ and $h_{\mu\nu}$ in mind *individually*, or whether should assess the background independence of the spin-2 theory with respect to the *composite* object $\hat{g}_{\mu\nu}+h_{\mu\nu}$, i.e. the metric $g_{\mu\nu}$ of general relativity (recall §4.1).[61] Bringing all this together, we have the following three cases:[62]

1. $\hat{g}_{\mu\nu}$ is regarded as geometrical; $h_{\mu\nu}$ is not regarded as geometrical. In this case, for each $\hat{g}_{\mu\nu}$ there are various distinct configurations of the $h_{\mu\nu}$ field, associated with the various solutions of the spin-2 theory for this background $\hat{g}_{\mu\nu}$. As such, geometrical and physical degrees of freedom do not coincide, and the theory counts as (fully) background dependent, for Belot.

2. $\hat{g}_{\mu\nu}$ is regarded as geometrical, as is $h_{\mu\nu}$, and background independence of the theory is assessed with respect to these objects, treated as fundamental. In this case, all the physical degrees of freedom are geometrical degrees of freedom, so the theory qualifies as background independent.

3. $\hat{g}_{\mu\nu}$ is regarded as geometrical, as is $h_{\mu\nu}$, and background independence of the theory is assessed with respect to the composite object $g_{\mu\nu}$, i.e. the metric of the associated model of general relativity. In this case, we again have it that all geometrical degrees of freedom are physical degrees of freedom, and so the theory comes out as fully background independent.

What we find here, then, is that only on option (1) does the spin-2 theory come out as background dependent, on Belot's account. In the next subsection, we'll see these verdicts mirrored when we turn to the background independence of SFT.

6.3 Background independence of SFT

So much for the warm-up; let's turn now to a systematic assessment of the background independence of SFTs. Our strategy for doing this will be as follows.

[60] It would be interesting to assess the status of $h_{\mu\nu}$ as spatiotemporal with respect to the criteria presented by Martens and Lehmkuhl (2020a,b); for our purposes here, though, we don't need to engage in that detailed study.

[61] These issues are discussed in the context of Belot's approach to background independence by Read (2023, ch. 3).

[62] In all of these cases, we're setting aside diffeomorphisms: we assume that diffeomorphisms always relate geometrically/physically equivalent models. Of course, this is somewhat controversial, in light of the points raised by Belot (2018), and further discussed by Read (2023). But we'll set aside the issues here.

First, in §6.3.1, we'll assess the background independence of each individual SFT (which, as we'll see, is akin to assessing the background independence of individual spin-2 theories as discussed in the previous subsection). Then, in §6.3.2, we'll assess the background independence of SFTs with respect to their invariant structure (which is akin to assessing the background independence of spin-2 theories with respect to their invariant general relativistic structure, as was also discussed in the previous subsection). And finally, in §6.3.3, we'll assess the background independence of what we were calling above (prior to philosophical analysis, of course) the 'manifestly background independent' formulations of these theories.

6.3.1 Individual theories

Our first task is to assess the background independence of an *individual* SFT, with models $\langle M, \hat{B}, \hat{\Psi} \rangle$ where $\hat{B}$ is the fixed background and $\hat{\Psi}$ is the string field (recall §3.5). As we'll see, the verdicts here mirror those which we gave for the understanding of the spin-2 theory which we presented in the previous subsection—i.e., on the understanding of spin-2 gravity according to which each background $\hat{g}_{\mu\nu}$ yields a *distinct* spin-2 theory.

All of this is quite straightforward to see. Since $\hat{B}$ is fixed (up to isomorphism) across models, it at least counts as an absolute object and absolute field—although perhaps not a fixed field, since it needn't be fixed *identically* in all kinematical possibilities. Still, in general one sees that an individual SFT violates 'absolute object-like' definitions of background independence, as one would expect.

Consider now the definition of background independence in terms of variational principles. The first issue to settle here is the action principle which one should be considering; this is the action given in (28). Recall that in that expression, we ignored ghost fields for simplicity, but following Erbin (2021, §10.4) they could be restored, and indeed for our purposes now it will be important that we consider them.

If one considers this action, then one sees that the background independence of an individual SFT appears to be violated, for our fixed, non-variational fields appear at low orders, and these fields are not subject to Hamilton's principle, from the spacetime perspective. In addition, one might think that this definition of background independence is violated doubly by virtue of the presence of the ghost fields; however, one has to be careful, because these fields *are* in fact varied (see Erbin (2021, §10.4)). Moreover, while presumably one wouldn't want to regard these as being 'physical' fields, even here there are delicate issues, as discussed by Dougherty (2021).[63] And as a final point on this variational approach to background independence, note that if one considers the SFT from the *worldsheet* point of view, then the fixed background fields are re-interpreted as coupling constants, and as such do not so obviously violate this definition of background independence (this point is discussed in the context of perturbative

[63] It would be nice to have some principled distinction between ghosts and Lagrange multiplier fields *vis-á-vis* their 'physicality'. But we won't pursue this here.

string theory by Read (2023, ch. 4)). All-in-all, then, while there are interesting subtleties here regarding ghosts, what ultimately one finds is that whether an individual SFT satisfies or violates the variational definition of background independence hinges crucially on whether one adopts (respectively) the worldsheet or spacetime point of view.

Finally, we have Belot's definition of background independence. Here, *prima facie* the case might seem relatively clear cut: since the background field $\hat{B}$—plausibly understood as the background spacetime geometry in the theory—is fixed from model to model (up to isomorphism), there are no geometrical degrees of freedom; in which case, an individual SFT is fully background dependent on Belot's account. This verdict would track case (1) in the previous subsection. That said, there are subtleties here again: if we take the background metric *together with* the metric fluctuation of the string field as being the geometrical degrees of freedom, then in fact geometrical degrees of freedom do vary from model to model. And—assuming that these co-vary with the non-geometrical degrees of freedom (encoded in the other excitations of the string field)—one will thereby arrive at the verdict that an individual SFT is fully background *independent*, in Belot's sense. This latter way of understanding the background independence of SFT in Belot's analysis is akin to case (2) in the previous subsection, save for the fact that we now also have material fields, encoded (as mentioned) in the other non-metrical excitations of the string field. (In addition, one might wish to regard the Kalb–Ramond field as being geometrical in string theory, since it mixes with the metric under T-duality: see Blumenhagen et al. (2013, §14.2).)

6.3.2 Invariant structure

So much for assessments of the background independence of individual SFTs, interpreted 'literally', where there is a majority (but hardly unanimous) verdict that such theories are background *dependent*. The next issue to consider is whether adjudications on the background independence of SFTs change when one interprets SFTs via their invariant (supergravity) structure, as discussed in §5.1 and §5.2. Here, we'll first consider open SFT, before turning to closed SFT.

Consider first our 'absolute object'-style definitions, with the invariant dynamical (super)gravity fields (89) in mind. Here, we see (again) that definition 1 ('no absolute objects') is violated, for the usual reason that (e.g.) the determinant of $G_{\mu\nu}$ will count as an absolute object. On the other hand, both definition 2 and definition 3 will be satisfied, since the spacetime fields (89) can vary (beyond isomorphism) from model to model; they are neither absolute fields nor fixed fields.

On the other hand, it is a little difficult to assess whether definition 4—in terms of variational principles—is satisfied when one is focusing on these invariant structures, the reason being that (unlike in the case of the spacetime fields in perturbative string theory as considered by Read (2023, ch. 5)) the action for these background (super)gravity fields is not known. Given this, it is difficult—absent further technical work—to ascertain the background independence of this

'invariant' structure according to the variational principles conception of background independence. Failing this, it seems that one might have to fall back on the action (28) for individual SFTs, which (as we've already seen) violates definition 4.

Finally, consider Belot's approach to background independence. With sensible choices of 'geometrical' and 'physical' fields (plausibly, $G_{\mu\nu}$ versus $B_{\mu\nu}$, Φ, etc.—though recall our above parenthetical remark on the Kalb–Ramond field), one sees that geometrical degrees of freedom do indeed match physical degrees of freedom, making the theory background independent. (This mirrors the verdicts on the spacetime formulation of perturbative string theory considered by Read (2023, ch. 5)).

When we turn now from the closed SFT to the open SFT, we of course still have a fixed closed string background; this background will end up violating definition 1 (absolute objects) and (plausibly) definition 4 (variational principles—to the extent that one can think about variational principles at all in the absence of a specific action). (At least on the assumption that there can't be non-isomorphic versions of this background in different models; if there in fact can be non-isomorphic versions of this background in different models, then definition 1 is in fact satisfied.) Assuming that one adopts the minor amendment to Belot's proposal for background independence made by Read (2023, ch. 3)—namely, that *every* object designated as 'geometrical' co-vary with the 'physical' degrees of freedom in one's models—these background closed string fields will also end up violating Belot's definition of background independence.

6.3.3 Manifestly background independent formulations

Finally, let's turn to what we called in table 1 the 'manifestly' background independent formulations of open and closed SFTs—that is, the formulations of open and closed SFTs described by the third and fourth column of that table. (Of course, the term 'manifestly' here should be taken with a pinch of salt, if one doesn't wish to cook the books in favour of a verdict of background independence—cf. Read (2023, ch. 5).)

Begin with Witten's BSFT for the open SFT. As before, by dint of being an open SFT, this theory has a closed SFT background in all models! (This is $\hat{C}$ in table 1.) Typically, such a background is not understood to be fixed identically in all models, and as such it does not lead to a violation of definition 3 of background independence (in terms of fixed fields). Whether or not it violates definition 1 (absolute objects), on the other hand, will depend upon whether one countenances models of BSFT with non-isomorphic versions of $\hat{C}$—if one does this, then (mirroring our above discussion of the spin-2 theory) this field will not violate this definition of background independence; otherwise, it will lead to such a violation (in our view, mirroring the case of individual spin-2 theories, there is a case to be made that it is sensible to regard each BSFT as being associated with a specific $\hat{C}$). In any case, since this background is not subject to Hamilton's principle in BSFT actions such as (95), it seems that Witten's BSFT *does* violate definition 4 of background independence (variational principles).

Finally, as before, whether or not Belot's definition of background independence is satisfied is again going to be a function of (a) whether different models of the BSFT can include non-isomorphic versions of $\hat{C}$, (b) whether other fields in BSFT are regarded as being 'geometrical' (although the grounds on which one would regard the $\mathcal{V}$ in BSFT as being 'geometrical' are not clear), and (c) whether one again adopts the above-mentioned proposed amendment to Belot's proposal made by Read (2019, ch. 3) (on which even if one were to identify other fields in BSFT as being 'geometrical', the theory would still not be background independent).

Turn now to the cZ action of the classical bosonic closed SFT presented by Ahmadain et al. (2024). In this case, one has an action (96) on the space of 2D CFTs which makes no reference to any fixed background whatsoever! As such, concerns regarding an outstanding fixed background which we saw for Witten's BSFT do not carry over to this case, and—*prima facie* (one has to keep in mind the limitations of the approach given that the construction works for tree-level only)—one appears to have secured an approach to SFT which on almost all philosophical accounts of the notion qualifies as background independent.[64]

Acknowledgments

We are extremely grateful to Subhroneel Chakrabarti for many helpful discussions and significant guidance over the course of this project. We're also indebted to Sebastian de Haro for detailed comments on a previous draft of this article. Finally, for valuable feedback we thank Nick Huggett, Chip Sebens, and audiences at Foundations of Physics 2025 (in Gdańsk) and EPSA 2025 (in Groningen).

References

Abouelsaood, A. et al. (1987). "Open strings in background gauge fields". *Nuclear Physics B* 280, pp. 599–624. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(87)90164-7. URL: https://www.sciencedirect.com/science/article/pii/0550321387901647.

Ahmadain, Amr, Frenkel, Alexander, and Wall, Aron C. (Oct. 2024). "A Background-Independent Closed String Action at Tree Level". arXiv: 2410.11938 [hep-th].

Anderson, J. L. (1967). *Principles of Relativity Physics*. Academic Press. ISBN: 9780120584505. URL: https://books.google.co.uk/books?id=mAVRAAAAMAAJ.

Belot, Gordon (2011). "Background-Independence". *General Relativity and Gravitation* 43, pp. 2865–2884.

— (2018). "Fifty Million Elvis Fans Can't be Wrong". *Noûs* 52.4, pp. 946–981.

[64] One cost of the generality of the approach of Ahmadain et al. (2024) is, of course, that the physical interpretation of (96) isn't obvious. But picking up this task is not our central concern here.

Bielińska, Marta and Jacobs, Caspar (2024). *A Philosophical Introduction to Hidden Symmetries in Physics*. URL: https://philsci-archive.pitt.edu/23922/.

Blumenhagen, Ralph, Lüst, Dieter, and Theisen, Stefan (2013). *Basic concepts of string theory*. Theoretical and Mathematical Physics. Heidelberg, Germany: Springer. ISBN: 978-3-642-29496-9. DOI: 10.1007/978-3-642-29497-6.

Brown, Harvey R. (2005). *Physical Relativity: Space-Time Structure From a Dynamical Perspective*. Oxford, GB: Oxford University Press UK.

Cattaneo, Alberto S., Mnev, Pavel, and Schiavina, Michele (2025). "BV Quantization". In: *Encyclopedia of Mathematical Physics (Second Edition)*. Ed. by Richard Szabo and Martin Bojowald. Oxford: Academic Press, pp. 543–555. ISBN: 978-0-323-95706-9. DOI: https://doi.org/10.1016/B978-0-323-95703-8.00095-1. URL: https://www.sciencedirect.com/science/article/pii/B9780323957038000951.

Chiaffrino, Christoph and Sachs, Ivo (June 2019). "Classical open string amplitudes from boundary string field theory". *Journal of High Energy Physics* 2019.6. ISSN: 1029-8479. DOI: 10.1007/jhep06(2019)086. URL: http://dx.doi.org/10.1007/JHEP06(2019)086.

Cho, Minjae, Collier, Scott, and Yin, Xi (2020). "Strings in Ramond-Ramond backgrounds from the Neveu-Schwarz-Ramond formalism". *Journal of High Energy Physics* 2020.12, p. 123. DOI: 10.1007/JHEP12(2020)123. URL: https://doi.org/10.1007/JHEP12(2020)123.

Dai, Jin, Leigh, R. G., and Polchinski, Joseph (2025/06/25 1989). "New connections between string theories". *Modern Physics Letters A* 04.21, pp. 2073–2083. DOI: 10.1142/S0217732389002331. URL: https://doi.org/10.1142/S0217732389002331.

Dolgachev, Igor (1999). "Introduction to string theory for mathematicians".

Dougherty, John (2021). "I Ain't Afraid of No Ghost". *Studies in History and Philosophy of Science Part A* 88.C, pp. 70–84. DOI: 10.1016/j.shpsa.2021.05.005.

Erbin, Harold (2021). *String Field Theory*. Springer International Publishing. DOI: 10.1007/978-3-030-65321-7. URL: https://doi.org/10.1007%2F978-3-030-65321-7.

— (2022). *Talk: Background Independence in String Theory*. URL: https://harolderbin.com/research-notes/.

Erler, Theodore (Oct. 2022). "Four lectures on analytic solutions in open string field theory". *Physics Reports* 980, pp. 1–95. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2022.06.004. URL: http://dx.doi.org/10.1016/j.physrep.2022.06.004.

Erler, Theodore and Maccaferri, Carlo (2014). "String field theory solution for any open string background". *Journal of High Energy Physics* 2014.10, p. 29. DOI: 10.1007/JHEP10(2014)029. URL: https://doi.org/10.1007/JHEP10(2014)029.

Erler, Theodore and Maccaferri, Carlo (2020). "String field theory solution for any open string background. Part II". *Journal of High Energy Physics*

2020.1, p. 21. DOI: 10.1007/JHEP01(2020)021. URL: https://doi.org/10.1007/JHEP01(2020)021.

Fletcher, Samuel C. (2020). "On Representational Capacities, with an Application to General Relativity". *Foundations of Physics* 50.4, pp. 228–249. DOI: 10.1007/s10701-018-0208-6.

Friedman, Michael (1983). *Foundations of Space-Time Theories: Relativistic Physics and the Philosophy of Science*. Princeton University Press.

Green, Michael B., Schwarz, John H., and Witten, Edward (2012). *Superstring Theory: 25th Anniversary Edition: Volume 1: Introduction*. Vol. 1. Cambridge Monographs on Mathematical Physics. Cambridge: Cambridge University Press. ISBN: 9781107029118. DOI: DOI:10.1017/CBO9781139248563. URL: https://www.cambridge.org/core/books/superstring-theory/35152940E2D9C3B0BD327EB7FF74DBBE.

Hořava, Petr (1989). "Background duality of open-string models". *Physics Letters B* 231.3, pp. 251–257. DOI: https://doi.org/10.1016/0370-2693(89)90209-8. URL: https://www.sciencedirect.com/science/article/pii/0370269389902098.

Huggett, Nick and Vistarini, Tiziana (2015). "Deriving General Relativity From String Theory". *Philosophy of Science* 82.5, pp. 1163–1174. DOI: 10.1086/683448.

Huggett, Nick and Wüthrich, Christian (2025). *Out of Nowhere*. Oxford University Press.

Hull, Chris (2025). *A New Action for Superstring Field Theory*. arXiv: 2508.03902 [hep-th]. URL: https://arxiv.org/abs/2508.03902.

Kajiura, Hiroshige (2002). "Homotopy algebra morphism and geometry of classical string field theories". *Nuclear Physics B* 630.3, pp. 361–432. DOI: https://doi.org/10.1016/S0550-3213(02)00174-8. URL: https://www.sciencedirect.com/science/article/pii/S0550321302001748.

Krátký, Matěj and Read, James (Dec. 2024). "On Functional Freedom and Penrose's Critiques of String Theory". *Philosophy of Physics*. DOI: 10.31389/pop.100.

Linnemann, Niels, Smeenk, Chris, and Baker, Mark Robert (Mar. 2023). "GR as a classical spin-2 theory?" *Philosophy of Science* 90.5, pp. 1363–1373. ISSN: 1539-767X. DOI: 10.1017/psa.2023.56. URL: http://dx.doi.org/10.1017/psa.2023.56.

Martens, Niels C. M. and Lehmkuhl, Dennis (2020a). "Cartography of the space of theories: An interpretational chart for fields that are both (dark) matter and spacetime". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 72, pp. 217–236. DOI: https://doi.org/10.1016/j.shpsb.2020.08.004. URL: https://www.sciencedirect.com/science/article/pii/S1355219820301106.

— (2020b). "Dark matter = modified gravity? Scrutinising the spacetime–matter distinction through the modified gravity/ dark matter lens". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 72, pp. 237–250. DOI: https://doi.org/10.1016/

j.shpsb.2020.08.003. URL: https://www.sciencedirect.com/science/article/pii/S135521982030109X.

Mazel, Ben, Wang, Charles, and Yin, Xi (2025). *Diffeomorphism in Closed String Field Theory*. arXiv: 2504.12290 [hep-th]. URL: https://arxiv.org/abs/2504.12290.

Møller-Nielsen, Thomas (2017). "Invariance, Interpretation, and Motivation". *Philosophy of Science* 84.5, pp. 1253–1264.

Muenster, Korbinian and Sachs, Ivo (2012). *Homotopy Classification of Bosonic String Field Theory*. arXiv: 1208.5626 [hep-th]. URL: https://arxiv.org/abs/1208.5626.

Nakahara, M. (2003). *Geometry, Topology and Physics, Second Edition*. Graduate student series in physics. Taylor & Francis. ISBN: 9780750306065. URL: https://books.google.co.uk/books?id=cH-XQB0Ex5wC.

Pitts, J. Brian (2006). "Absolute Objects and Counterexamples: Jones-Geroch Dust, Torretti Constant Curvature, Tetrad-Spinor, and Scalar Density". *Studies in History and Philosophy of Modern Physics* 37.2, pp. 347–371.

Polchinski, Joseph (1998). *String Theory: Volume 1: An Introduction to the Bosonic String*. Vol. 1. Cambridge Monographs on Mathematical Physics. Cambridge: Cambridge University Press. ISBN: 9780521633031. DOI: DOI:10.1017/CBO9780511816079. URL: https://www.cambridge.org/core/books/string-theory/30409AF2BDE27D53E275FDA395AB667A.

Pooley, Oliver (2017). "Background Independence, Diffeomorphism Invariance, and the Meaning of Coordinates". In: *Towards a Theory of Spacetime Theories*. Ed. by Dennis Lehmkuhl, Gregor Schiemann, and Erhard Scholz. Birkhäuser.

Read, James (2019). "On Miracles and Spacetime". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 65, pp. 103–111. DOI: 10.1016/j.shpsb.2018.10.002.

— (2023). *Background independence in classical and quantum gravity*. Oxford University Press.

— (2025). "Good VIBES only". *Synthese*.

Read, James and Møller-Nielsen, Thomas (2020). "Motivating Dualities". *Synthese* 197, pp. 263–291.

Salimkhani, Kian (2020). "The Dynamical Approach to Spin-2 Gravity". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 72, pp. 29–45. DOI: 10.1016/j.shpsb.2020.05.002.

Sen, Ashoke (2018). "Background independence of closed superstring field theory". *Journal of High Energy Physics* 2018.155. DOI: 10.1007/JHEP02(2018)155. URL: https://doi.org/10.1007/JHEP02(2018)155.

Sen, Ashoke and Zwiebach, Barton (1994a). "A proof of local background independence of classical closed string field theory". *Nuclear Physics B* 414.3, pp. 649–711. DOI: https://doi.org/10.1016/0550-3213(94)90258-5. URL: https://www.sciencedirect.com/science/article/pii/0550321394902585.

— (1994b). "Quantum background independence of closed-string field theory". *Nuclear Physics B* 423.2, pp. 580–630. DOI: https://doi.org/10.1016/

0550-3213(94)90145-7. URL: https://www.sciencedirect.com/science/article/pii/0550321394901457.

Sen, Ashoke and Zwiebach, Barton (2024). *String Field Theory: A Review*. arXiv: 2405.19421 [hep-th]. URL: https://arxiv.org/abs/2405.19421.

Stachel, John (2007). "The Story of Newstein or: Is Gravity Just Another Pretty Force?" In: *The Genesis of General Relativity*. Ed. by Michel Janssen et al. Dordrecht: Springer Netherlands, pp. 1962–2000. ISBN: 978-1-4020-4000-9. DOI: 10.1007/978-1-4020-4000-9{_}46. URL: https://doi.org/10.1007/978-1-4020-4000-9_46.

Totsuka-Yoshinaka, Jojiro (2025). *BSFT-like action from cohomomorphism*. arXiv: 2504.10061 [hep-th]. URL: https://arxiv.org/abs/2504.10061.

Weatherall, James Owen (2016). "Understanding Gauge". *Philosophy of Science* 85 (5), pp. 1039–1049.

Witten, Edward (1986). "Non-commutative geometry and string field theory". *Nuclear Physics B* 268.2, pp. 253–294. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(86)90155-0. URL: https://www.sciencedirect.com/science/article/pii/0550321386901550.

— (Dec. 1992). "On background-independent open-string field theory". *Phys. Rev. D* 46 (12), pp. 5467–5473. DOI: 10.1103/PhysRevD.46.5467. URL: https://link.aps.org/doi/10.1103/PhysRevD.46.5467.