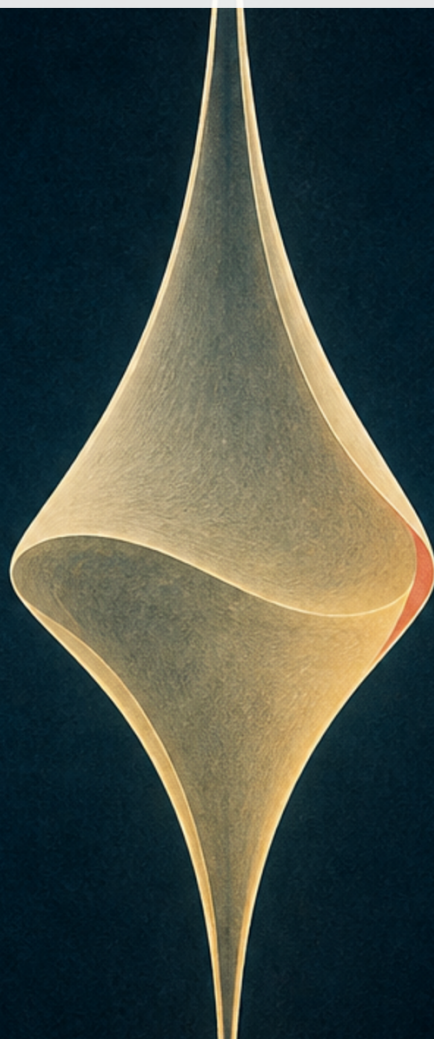


String Theory and M-Theory

Supplementary track, Fall 2010



Lectures by Leonard Susskind

Companion edition by [LazyingArt LLC](#) with [Video2Book](#)

About These Notes

These independently edited companion notes follow lectures by Leonard Susskind. They were reconstructed from machine transcripts, subtitles, and selected blackboard frames, then checked against the available source material. They are not an original manuscript by Professor Susskind and have not been reviewed or endorsed by him.

The edition was prepared by [LazyingArt LLC](#) using the open-source [Video2Book](#) workflow. Machine transcripts remain available separately and may contain recognition errors. Editorial clarifications and nontrivial reconstructions are identified in footnotes.

Contents

1	From Hadronic Regge Trajectories to Light-Cone Strings	1
1.1	The hadronic origin of string theory	1
1.2	Duality in pion–pion scattering	2
1.3	One world sheet, two slicings	4
1.4	Flux tubes and the change of scale	7
1.5	Two expansions of relativistic energy	9
1.6	The infinite-momentum frame	13
1.7	An open string made from beads and springs	14
1.8	Why the Hamiltonian measures mass squared	17
1.9	The same object in two frames	20
1.10	A superconducting model of confinement	21
1.11	Transverse degrees of freedom and the size of a particle	23
2	From Boundary Conditions to the Open-String Spectrum	26
2.1	Calculus on the open-string interval	26
2.1.1	Dirichlet and Neumann endpoints	27
2.1.2	The Fourier bases	29
2.1.3	One oscillator formula to remember	30
2.2	What makes a composite system a particle?	31
2.3	Exact transverse mechanics at large momentum	33
2.3.1	Subtracting the boost and rescaling time	34
2.4	A discrete open string	36
2.5	The continuum wave system	38
2.6	Why a free endpoint is Neumann	40
2.7	Normal modes of the open string	42
2.7.1	Kinetic term and the zero mode	42

2.7.2	Stretching term and oscillator frequencies	43
2.8	Quantized levels and particle behavior	46
2.9	Why open-string interactions produce closed strings	48
3	From Open-String Quanta to Tachyons and Gravity	50
3.1	The oscillator hidden in each string mode	50
3.1.1	Two transverse towers	53
3.2	Massive spin and massless helicity	55
3.2.1	Linear and circular photon polarization	56
3.3	The lowest open-string levels	58
3.4	What a tachyon really means	61
3.4.1	Negative curvature of the field potential	62
3.4.2	The upright-pendulum chain	63
3.5	How strings interact	65
3.5.1	The open-string vertex	66
3.5.2	An open string can close itself	66
3.5.3	Closed strings can interact without opening	67
3.6	Why gravity cannot be avoided	68
3.7	The scale and coupling still missing	70
4	Closed Strings, Level Matching, and Duality	72
4.1	A Generator Kept in Reserve	72
4.2	A Coordinate Around a Loop	73
4.3	Periodicity and Running Waves	74
4.3.1	Complex Fourier modes	75
4.4	Energy in the Two Directions	76
4.4.1	Twice as many oscillators	78
4.5	The Spectrum Guess That Fails	79
4.5.1	Level matching, first as a rule	79
4.6	A Graviton in the First Matched Level	80
4.7	Why the Levels Must Match	82
4.7.1	The discrete version	82
4.7.2	The continuum version	83
4.7.3	The generator is the energy difference	85
4.8	Two Meanings of Left and Right	86

4.9	What Does Fundamental Mean?	87
4.9.1	The Wigner question	88
4.10	Toward M-Theory	89
5	Scales, Strings, and Twenty-Six Dimensions	90
5.1	The Scale We Had Left Out	90
5.1.1	Dimensions and oscillator gaps	92
5.2	Units Made from Universal Constants	93
5.2.1	The Planck area first	94
5.3	How Close Should the Two Scales Be?	97
5.3.1	What running couplings can tell us	97
5.4	Why Direct String Spectroscopy Is Remote	99
5.4.1	Extra dimensions and the localization limit	100
5.5	The Ground State Must Sit Below the Photon	101
5.6	Subtract the Motion Before Reading the Spectrum	102
5.7	Regulating the Sum	104
5.8	Counting the Transverse Directions	110
5.8.1	The closed string checks the count	111
6	Scattering, Duality, and the Veneziano Amplitude	112
6.1	What a Theory Must Predict	112
6.2	Four-Momentum and the All-Incoming Convention	113
6.3	Mandelstam Variables	115
6.4	Direct and Crossed Channels	117
6.5	One Function, Two Channel Descriptions	118
6.6	A Strip Swept Out by an Open String	119
6.7	The Oscillator Calculation in Slow Motion	120
6.8	From Propagation Time to the Beta Function	122
6.9	What the Scattering Test Established	125
7	Worldsheets, Wick Rotation, and Channel Duality	127
7.1	From a Worldline to a Sum over Histories	127
7.2	Turning Time Sideways	129
7.3	From Worldlines to Worldsheets	130
7.4	The Euclidean Equation of Motion	133

7.5	Conformal Coordinates	136
7.6	The Slit Sheet and the Disc	138
7.7	Two Limits of One Surface	140
8	Conformal Maps and the Open-String Scattering Prescription	143
8.1	Electrostatics in a Two-Dimensional World	143
8.2	A Derivative with No Preferred Direction	145
8.3	Harmonic Conjugates and Conformal Geometry	146
8.3.1	Why Angles Are Preserved	147
8.4	Examples that Train the Eye	148
8.4.1	Squaring and Complex Conjugation	148
8.4.2	The Exponential and Its Inverse	149
8.5	Half-Plane, Strip, Cylinder, and Disc	150
8.5.1	The Logarithm Turns Radius into Time	150
8.5.2	A Cayley Map to the Unit Disc	152
8.6	The Four-String Worldsheet	154
8.7	External States as Boundary Insertions	156
8.8	The Gaussian Integral as Electrostatics	159
8.8.1	Conformal Redundancy and the Number of Moduli	161
8.9	What the Prescription Does Not Yet Solve	162
9	Strings, Gravity, Compactification, and T-Duality	164
9.1	Consistency as a Selection Principle	164
9.2	The Point Particle on a Sphere	165
9.3	The Vacuum String Is Not Pointlike	166
9.4	When Internal Spreading Changes the Geometry	169
9.5	The Metric Must Reach a Fixed Point	171
9.5.1	The Same Condition Is the Vacuum Einstein Equation	173
9.6	Compact Dimensions Rather Than Missing Dimensions	174
9.7	The Shape Parameters of a Flat Torus	178
9.8	Reading a Compact Circle from Its Spectrum	179
9.9	Winding States Complete the Spectrum	181
9.10	T-Duality: Small and Large Circles Are the Same Physics	183
9.11	Why Winding Sectors Cannot Be Omitted	184
9.12	Can One Hear the Shape of a String Background?	186

10 T-Duality, D-Branes, and Gauge Theory	188
10.1 Why Return to T-Duality?	188
10.2 Winding Is Conserved Even When Strings Reconnect	190
10.3 Momentum Modes and Winding Modes	191
10.4 The Worldsheet Form of the Duality	194
10.5 Why Winding States Cannot Be Avoided	195
10.6 Momentum Charge from Five-Dimensional Gravity	196
10.7 The Second Vector Field and Winding Charge	198
10.8 Open Strings Change the Question	200
10.9 The Object Holding the Endpoint Is a D-Brane	202
10.10 Counting the Dimensions of a D-Brane	203
10.11 Open Strings Become Fields on a Brane	206
10.12 A Stack of Branes Produces Nonabelian Color	206
10.13 D-Strings End as Magnetic Monopoles	209
10.14 What the Construction Does and Does Not Claim	211

Chapter 1

From Hadronic Regge Trajectories to Light-Cone Strings

Overview. String theory did not begin as a theory of quantum gravity. It began as an attempt to understand the spectrum and scattering of hadrons. Straight Regge trajectories suggested an extended object; duality in pion–pion scattering suggested that two apparently distinct processes were two slicings of one surface. We shall follow those clues into the world-sheet picture, reinterpret the hadronic string as a QCD flux tube, and then use the infinite-momentum frame to derive the transverse mechanics of an open string. The central result is that its effective internal Hamiltonian measures mass squared. A Hooke-law energy proportional to L^2 in that description therefore becomes a rest energy proportional to L , the defining behavior of a string with constant tension.

1.1 The hadronic origin of string theory

Around 1968–1970, the problem that led to strings concerned protons, neutrons, and especially mesons. The quark model was already a serious idea, but the quark content of hadrons had not yet been experimentally established in the modern sense. Gluons were even less prominent in ordinary discussion. Nambu had proposed related gauge-field ideas, but the settled language of quantum chromodynamics did not yet exist. The raw material was instead an unexpectedly rich spectrum of strongly interacting particles.

It is useful to ignore the proton–neutron distinction temporarily and call both states the nucleon. Beyond the nucleon were heavier states with larger spin, followed by still heavier states with still larger spin. Chew and Frautschi organized such families on a plot with angular momentum on the vertical axis and mass squared on the horizontal axis. In the historical units of the period the proton mass was about one BeV; the same unit is now called a GeV.

Measure angular momentum in units of $\hbar$. For a baryon family, the dimensionless spin values are half-integral,

$$J = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots, \quad (1.1)$$

while meson families have integral spin. The remarkable empirical fact was that either family lay approximately on a straight line when plotted against M^2 . In modern notation one writes

$$J \simeq \alpha_0 + \alpha' M^2, \quad (1.2)$$

where α_0 is an intercept and α' is the Regge slope.¹

The original straight line was not compelling because it passed through only two known points. Any two points determine a line. The surprise came later: newly discovered states continued to land close to the same line. Meson trajectories showed the same organization. The pion begins near $J = 0$ with a small mass on the GeV scale; the rho begins at $J = 1$; further excitations continue upward. Five, six, or seven states could appear in a family. Even more strikingly, different baryon and meson trajectories were nearly parallel. Thus an increase by one quantum of angular momentum corresponded to approximately the same increase in mass squared:

$$\Delta J = 1 \implies \Delta M^2 \simeq \frac{1}{\alpha'}. \quad (1.3)$$

Question from the audience. What does it mean for different Regge trajectories to be parallel?^a

Response. Choose one family of particles and plot all of its excited states; they lie along one line. Repeat this for another meson or baryon family. The lines have approximately the same slope even though their intercepts differ. The same increment of angular momentum therefore costs approximately the same increment of M^2 . This common slope was called the universal Regge slope.

^aLecture timestamp: 00:09:32.

The comparison with ordinary extended objects shows why the pattern mattered. A basketball acquires rotational energy as it is spun, but its trajectory is curved rather than linear, and at sufficiently high angular velocity the ball is torn apart. An atom also has rotational excitations, but eventually it is ionized; its trajectory ends. Hadrons displayed a much simpler sequence and, within the available data, no sign that the sequence was giving out.

This spectrum said that a hadron was not behaving like a point. A point has no spatial structure to spin or deform. An electron has no observed tower of higher-spin internal states and is correspondingly pointlike within present experimental bounds. A hadron, by contrast, can be spun up. Its spectrum suggests compositeness, and the relation between spin and mass hints that the object deforms as its angular momentum increases.

1.2 Duality in pion–pion scattering

Spectroscopy supplied one clue. Pion–pion scattering supplied another. Two pions can combine into a rho-like resonance and later emerge again as two pions. In a conventional Feynman diagram this is an annihilation and re-creation process, customarily called an s -channel process. Turning the same elementary diagram sideways gives a crossed, or t -channel, process in which a rho-like particle is exchanged between the pions.

¹Editorial clarification: The lecture emphasizes the observed linear relation and its nearly universal slope. The intercept notation is supplied only to state that relation compactly; the trajectories are empirical and approximate.

The rho is not alone. It belongs to a Regge family, so consistency appears to require the rho and all of its excited partners in either channel. Schematically,

$$\mathcal{A}(s, t) \simeq \sum_n \mathcal{A}_{\rho_n}^{(s)}(s, t), \quad \mathcal{A}(s, t) \simeq \sum_n \mathcal{A}_{\rho_n}^{(t)}(s, t). \quad (1.4)$$

These expressions are not two approximations that should simply be added. Numerical studies found that summing the tower of resonances in the annihilation channel already gave approximately the contribution expected from the exchange channel. Conversely, the exchanged tower could reproduce the resonance description. Adding both complete towers therefore appeared to count the same physics twice. Ordinary perturbative bookkeeping says that if two diagrams exist, both must be included. The hadronic amplitudes instead suggested that the two channel descriptions were dual.



Figure 1.1: The board joins the Chew–Frautschi plot to the annihilation and crossed-channel diagrams, making the spectral and scattering clues part of one problem.

Lecture frame: 00:14:20.

The same relations can be separated into a clean schematic. The first panel records the linear spectrum; the other two represent the alternative channel expansions of the same amplitude.

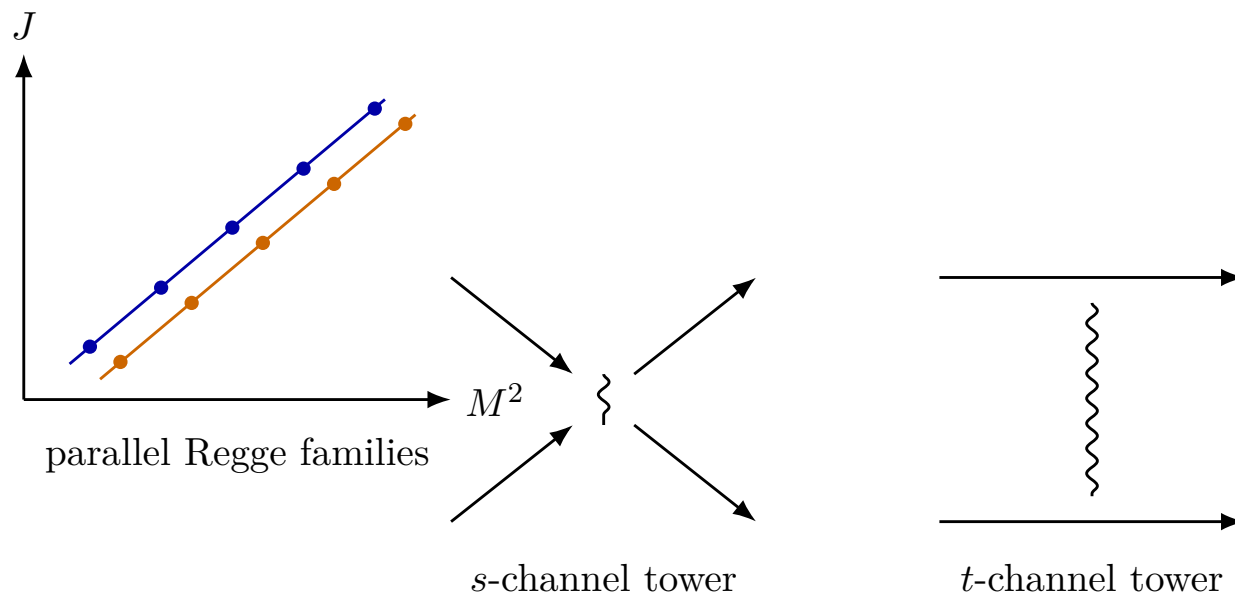


Figure 1.2: Parallel Regge families and the two channel expansions involved in hadronic duality.

1.3 One world sheet, two slicings

The natural response was to seek one geometrical history that could be read in either way. Represent a pion by a quark–antiquark pair and let time run upward. Instead of drawing only one-dimensional particle lines, fill in the region bounded by the quark trajectories. A horizontal slice through this two-dimensional history may show two incoming strings joining and later separating. A different slicing displays a string exchanged from one side to the other. The two Feynman-channel pictures are no longer independent mechanisms; they are different decompositions of one surface.

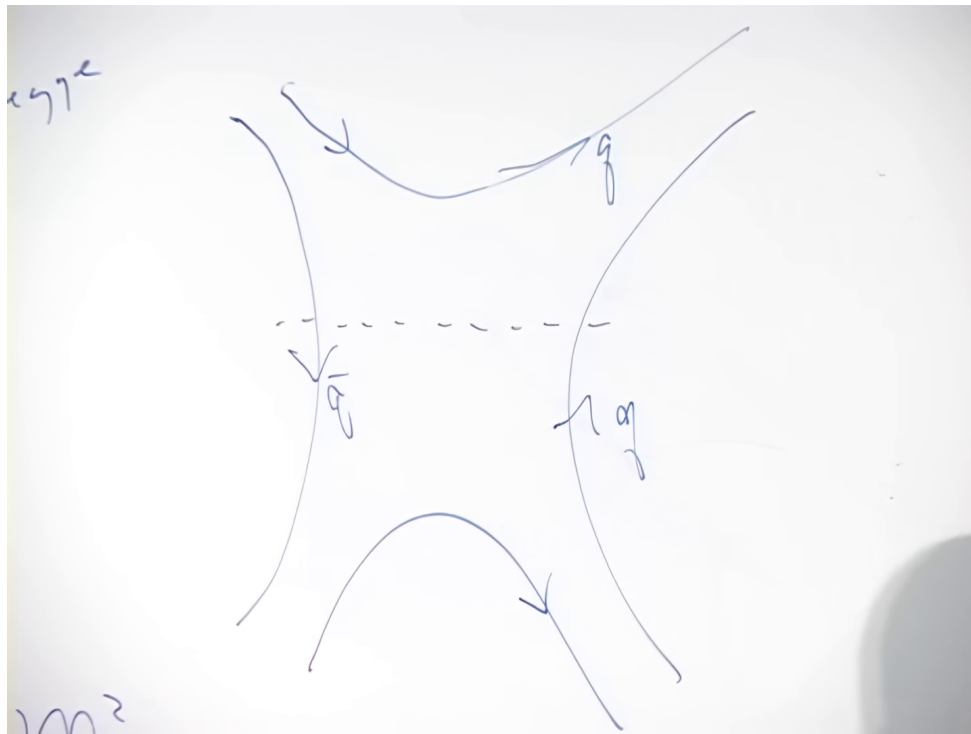


Figure 1.3: The quark-line topology on the board. Different cuts through the same two-dimensional history produce the resonance and exchange readings.

Lecture frame: 00:17:50.

The region between a quark and an antiquark supplies one further idea. At a fixed time, the slice contains two endpoints and something filling the whole interval between them. That object is one-dimensional: a string. During an interaction the string sweeps out the two-dimensional surface.

Definition 1.1. The spacetime history of a point particle is a *world line*. The spacetime history of a string is a *world sheet*. A string at one instant is a spatial slice of its world sheet.

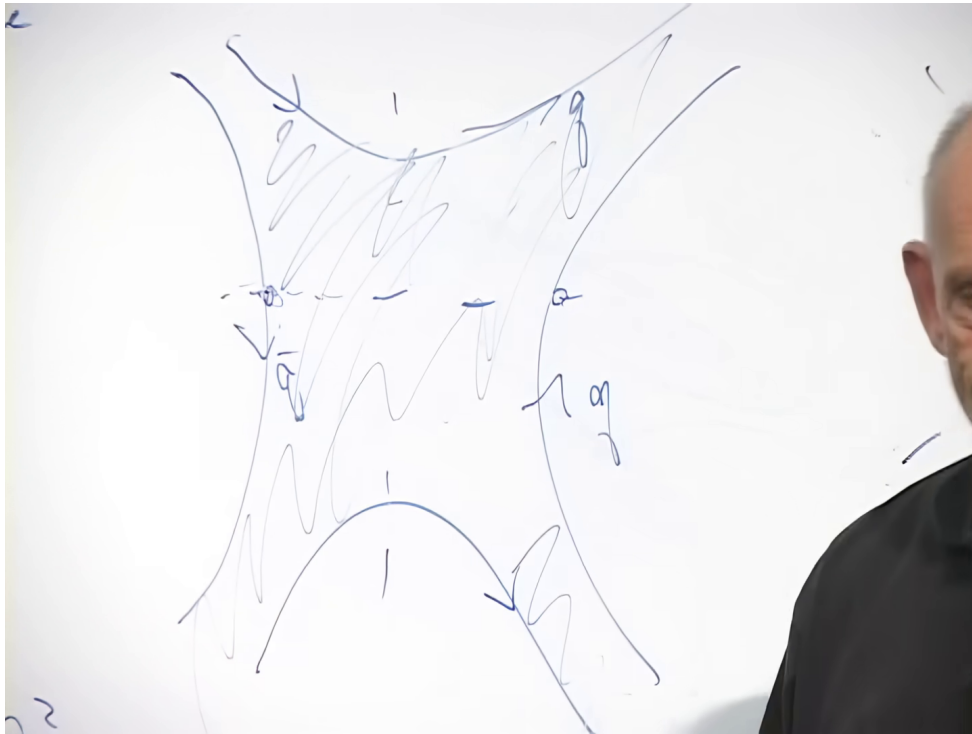


Figure 1.4: The board surface filled in between the endpoint trajectories. A time slice intersects it in a one-dimensional string.

Lecture frame: 00:19:15.

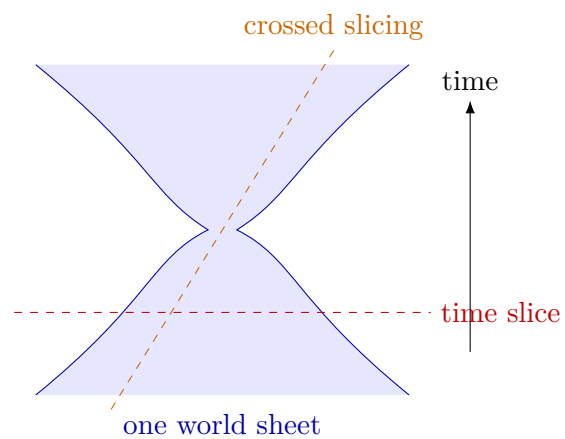


Figure 1.5: A clean world-sheet schematic. Distinct slicings organize the same surface into different intermediate-channel descriptions.

This geometric reconstruction was exploratory, not a deduction of the full theory from first principles. Different people reached the early string picture in different ways. Its strength was that the same extended object could address both observations: a string can rotate and deform, and its world sheet can encode the dual scattering channels.

Question from the audience. Why introduce a string-like bridge rather than an ordinary particle exchanged between the quarks?^a

Response. At one instant there is not merely a localized exchanged particle. There is a quark, an antiquark, and structure at all points between them. That structure might be fundamentally continuous or might be a dense collection of constituents approximating a continuum; the first picture did not decide. What mattered was that the whole interval was occupied and that the observed trajectories showed no tendency to terminate like an ordinary shoelace with weights on its ends.

^aLecture timestamp: 00:21:20.

Question from the audience. Can the Standard Model describe pion scattering? ^a

Response. Yes, although usefulness depends on the energy regime. At low energy the Standard Model, through QCD and its low-energy descriptions, accounts for a great deal of pion scattering. At higher energies, where whole towers of hadronic exchanges matter, direct calculations become more difficult and a string-like organization can be more useful.

^aLecture timestamp: 00:22:45.

Question from the audience. Did the early experiments establish that a Regge trajectory never ends?^a

Response. No. They established only that the observed sequence did not yet show an endpoint. The inference of indefinite stretchability was a suggestive extrapolation, not a measured infinity.

^aLecture timestamp: 00:23:37.

1.4 Flux tubes and the change of scale

Modern QCD gives the old picture a field-theoretic interpretation. A quark and an antiquark act as opposite sources of the gluon field. In ordinary electrodynamics, electric flux spreads sideways as opposite charges are separated. The field weakens with distance, so the region between the charges does not retain a fixed transverse profile. QCD is nonlinear. Its dynamics can confine the relevant chromoelectric flux to a narrow tube joining the quark and antiquark.

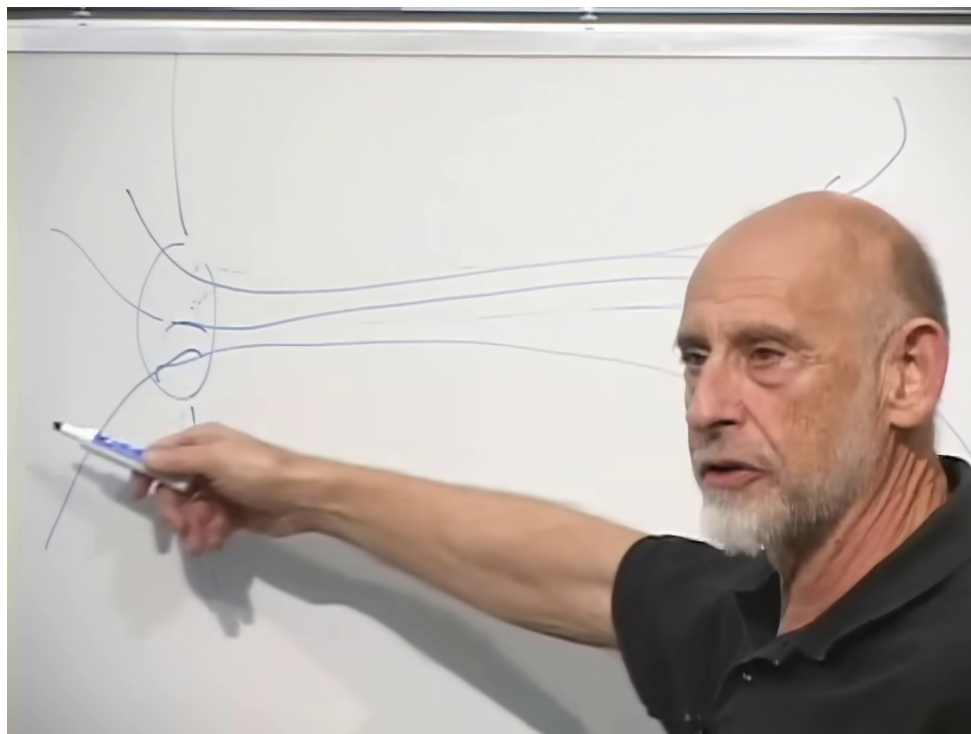


Figure 1.6: Quark and antiquark field lines drawn as a narrow flux tube rather than a freely spreading Coulomb field.

Lecture frame: 00:25:15.

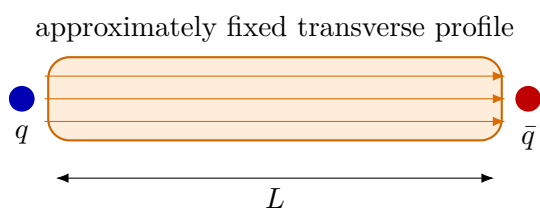


Figure 1.7: A confined flux tube lengthens without spreading. Fixed energy per unit length gives $E = TL$.

The particle language and field language say the same thing differently. In the field description, the tube retains roughly the same field strength and cross-sectional area as its endpoints separate. In a constituent picture, the stretching energy creates more gluonic structure between the endpoints. An ordinary rubber band has a fixed number of molecules, so its molecules separate until the material breaks. Imagine instead that a new molecule were inserted whenever a gap opened. The density along the band could remain fixed while the band became longer. That is the intended analogy for the gluon field: the tube gains length rather than thinning toward rupture.

Question from the audience. Would a string that made up an electron also be composed of gluons?^a

Response. No such conclusion follows. The hadronic string is a QCD object on the scale of a proton. A hypothetical fundamental string relevant to quantum gravity would live near the vastly smaller Planck scale, and there is no evidence here that an electron is a hadronic gluon string. The striking fact is that related mathematics arose at both scales, not that the physical objects are identical.

^aLecture timestamp: 00:28:08.

Attempts to turn the hadronic picture into a complete interacting string theory were promising but did not quite produce the desired theory of hadrons. The mathematical model repeatedly contained a massless particle of spin two. In a hadronic theory that state looked like an unwanted prediction; as a theory of gravity it had exactly the quantum numbers of the graviton. Joel Scherk and John Schwarz made the decisive reinterpretation: perhaps the theory was telling us about quantum gravity rather than hadrons. The hadronic and fundamental applications are physically different, even though the early hadron problem supplied the route to the mathematics.

Question from the audience. How do the flux lines go from a positive endpoint to a negative endpoint?^a

Response. They leave one source and enter the other, just as magnetic field lines run from the north pole to the south pole of a bar magnet. The opposite endpoint labels express the orientation of the flux; a bar magnet likewise has two opposite poles rather than two north poles.

^aLecture timestamp: 00:31:46.

Question from the audience. Is the string continuous or discrete? ^a

Response. Quantum theory makes that distinction subtle. The electromagnetic field is represented as a continuous field and also has discrete quanta. In the same way, a string may be treated continuously while its excitations are quantized. Both descriptions can be valid, depending on the question being asked.

^aLecture timestamp: 00:32:45.

Question from the audience. If the angle of the spacetime slice through a world sheet is changed, is the intersection still a string?^a

Response. Yes. Changing the simultaneity slice is a Lorentz transformation. The new slice intersects the surface in a string seen in a different state of motion. Even a string whose center of mass is at rest is not internally motionless; it may vibrate. “At rest” refers only to its center of mass.

^aLecture timestamp: 00:34:20.

1.5 Two expansions of relativistic energy

A direct relativistic description of an extended, vibrating object is awkward. The useful strategy is to find a frame in which its relevant motion has the form of ordinary nonrelativistic mechanics. Begin with a single particle. Nonrelativistically one writes

$$E_{\text{nr}} = \frac{\mathbf{p}^2}{2m} + B, \quad (1.5)$$

where B is an internal or binding energy independent of the overall momentum. For several noninteracting particles, both the kinetic terms and the internal constants are added. Since only energy differences govern the motion, a state-independent additive constant can often be suppressed.

The exact relativistic energy is

$$E = \sqrt{\mathbf{p}^2 c^2 + m^2 c^4}. \quad (1.6)$$

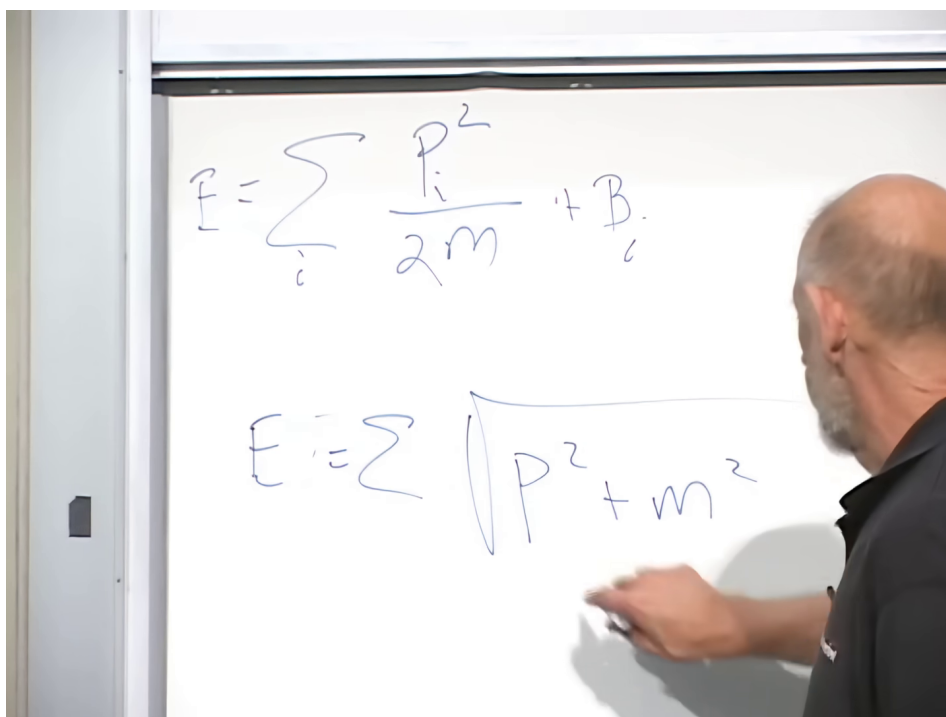


Figure 1.8: The nonrelativistic kinetic energy and the exact relativistic square root, with the powers of c corrected during the discussion.

Lecture frame: 00:38:45.

Question from the audience. Should the relativistic formula contain $p^2 c^2$ and $m^2 c^4$, rather than the powers initially written on the board? ^a

Response. Yes. Dimensional consistency fixes the expression to $E^2 = p^2 c^2 + m^2 c^4$. After recovering the nonrelativistic limit we shall set $c = 1$, but keeping the powers here makes the expansion transparent.

^aLecture timestamp: 00:39:12.

When $p^2 c^2 \ll m^2 c^4$, factor out $m^2 c^4$ and use $\sqrt{1 + \epsilon} \simeq 1 + \epsilon/2$:

$$E = mc^2 \sqrt{1 + \frac{p^2}{m^2 c^2}} \simeq mc^2 + \frac{p^2}{2m}. \quad (1.7)$$

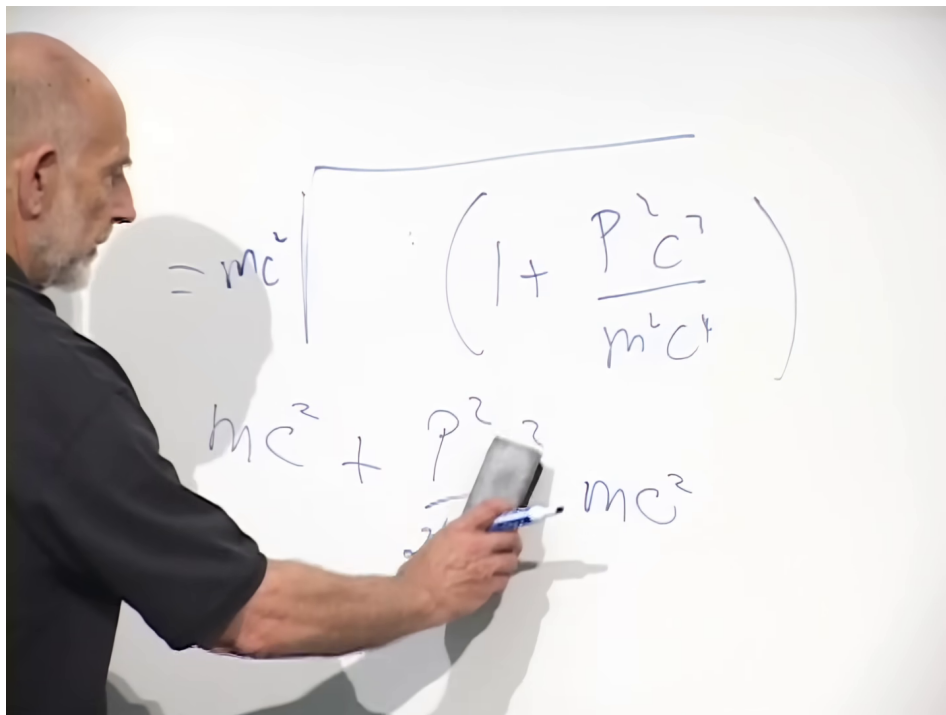


Figure 1.9: The small-momentum expansion recovering rest energy plus the familiar nonrelativistic kinetic term.

Lecture frame: 00:41:30.

This approximation requires more than a slowly moving center of mass. A box may drift slowly while particles inside it move near the speed of light. In that case the internal dynamics are still relativistic. The ordinary nonrelativistic limit requires all relevant constituent velocities to be small in the chosen frame.

There is another useful limit. Boost the entire system along the z -axis until every constituent has a large positive longitudinal momentum P_{zi} . Any constituent initially moving backward is carried forward by a sufficiently large boost. The transverse momenta and rest masses remain fixed:

$$P_{zi} \gg |p_{xi}|, |p_{yi}|, m_i, \quad p_{\perp i}^2 = p_{xi}^2 + p_{yi}^2, \quad (1.8)$$

where from this point onward $c = 1$. Expanding for large P_{zi} gives

$$\begin{aligned}
 E_i &= \sqrt{P_{zi}^2 + p_{\perp i}^2 + m_i^2} \\
 &= P_{zi} \sqrt{1 + \frac{p_{\perp i}^2 + m_i^2}{P_{zi}^2}} \\
 &\simeq P_{zi} + \frac{p_{\perp i}^2 + m_i^2}{2P_{zi}}.
 \end{aligned} \tag{1.9}$$

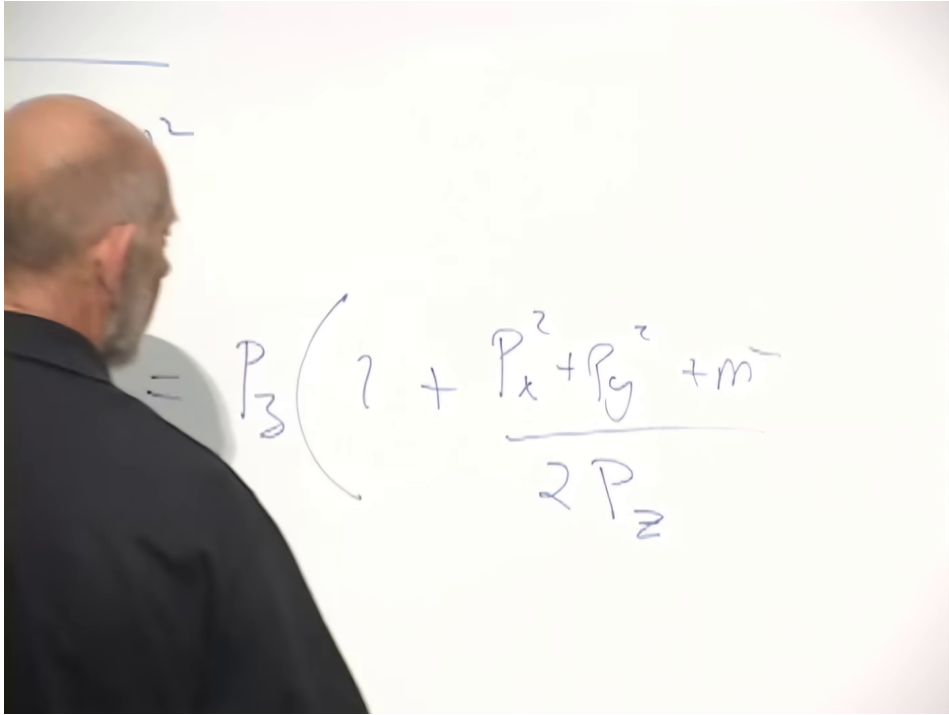


Figure 1.10: The large- P_z square-root expansion, with transverse momentum and mass retained as the small quantities.

Lecture frame: 00:46:45.

For many constituents,

$$E \simeq \sum_i P_{zi} + \sum_i \frac{p_{\perp i}^2}{2P_{zi}} + \sum_i \frac{m_i^2}{2P_{zi}}. \tag{1.10}$$

The first term is the conserved total longitudinal momentum. It is enormous, but it is also a fixed additive contribution. Keeping or subtracting it changes no energy difference, just as adding any other conserved number to a Hamiltonian would not change the equations of motion. The dynamical part is therefore

$$\begin{aligned}
 H_{\perp} &= E - P_z \simeq \sum_i \left(\frac{p_{\perp i}^2}{2P_{zi}} + \frac{m_i^2}{2P_{zi}} \right), \\
 P_z &= \sum_i P_{zi}.
 \end{aligned} \tag{1.11}$$

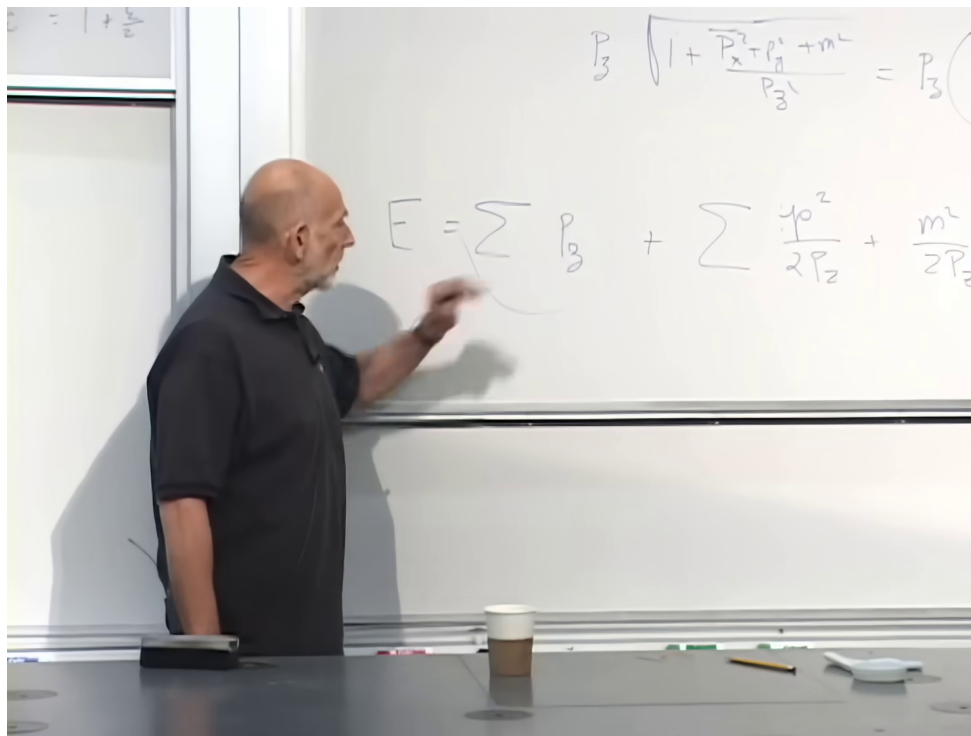


Figure 1.11: The many-particle decomposition into conserved longitudinal momentum, transverse kinetic energy, and mass-dependent internal terms.

Lecture frame: 00:49:16.

1.6 The infinite-momentum frame

The preceding limit was historically called the infinite-momentum frame and is now usually formulated in light-cone coordinates. Its meaning can be read directly from the Hamiltonian. In quantum mechanics,

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = H |\psi\rangle. \quad (1.12)$$

All nontrivial terms in $H_\perp$ scale as $1/P_z$. As the boost grows, energy splittings shrink and changes occur more slowly. This is time dilation. A radioactive system appears to decay more slowly when highly boosted; one can compensate by observing it for a proportionally longer time. The slowing does not remove the dynamics. It rescales their clock.

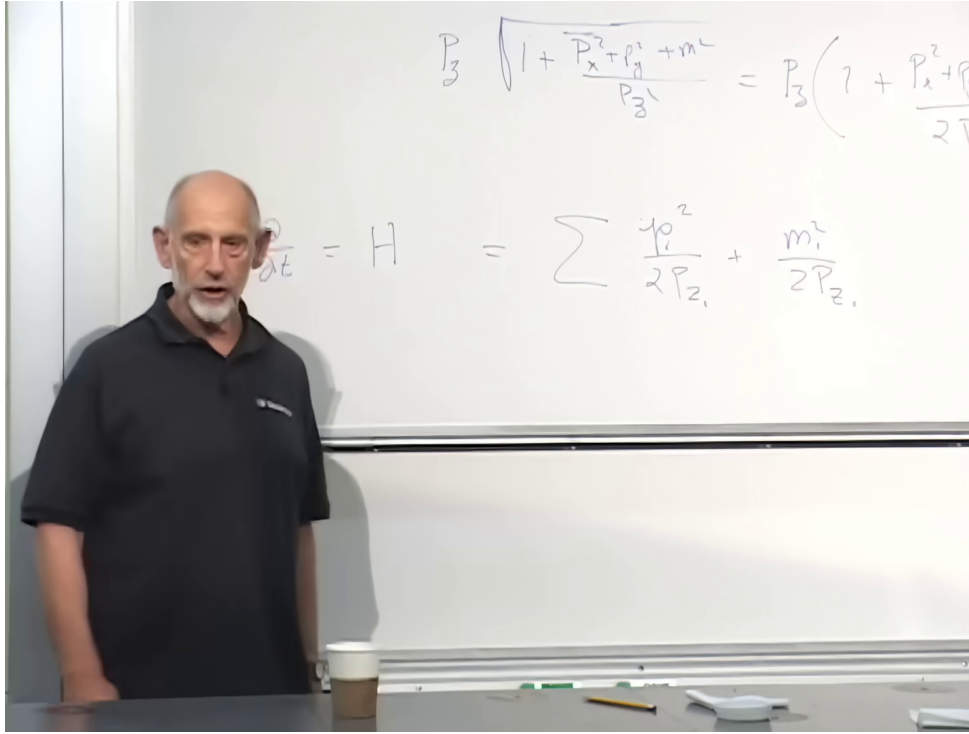


Figure 1.12: The board identifies the reduced energy with the Hamiltonian that generates the slowed time evolution of the boosted system.

Lecture frame: 00:52:57.

The transverse term has precisely the nonrelativistic form $p^2/(2m)$, but the role of mass is played by P_z . This is physically sensible: a larger longitudinal momentum makes a particle harder to deflect sideways, so it acts as transverse inertia. The term $m_i^2/(2P_{zi})$ does not depend on the two-dimensional state of motion and therefore plays the role of an internal or binding energy. Relativistic motion in three spatial dimensions has become nonrelativistic-looking motion in the two dimensions perpendicular to the boost.

This is the permission needed to use an elementary masses-and-springs model for a relativistic string. The model is not pretending that the string is slow in its rest frame. It uses the frame in which the transverse dynamics have the required nonrelativistic structure.

1.7 An open string made from beads and springs

Take an open string, meaning a string with two endpoints. Something may later be attached to those endpoints, but the present concern is the string itself. Approximate it by beads at transverse positions $\mathbf{X}_i(t) = (X_i(t), Y_i(t))$, joined by identical springs. A discrete Hamiltonian is

$$H_N = \sum_{i=1}^N \left[\frac{\mu_N}{2} \dot{\mathbf{X}}_i^2 + \frac{k_N}{2} (\mathbf{X}_{i+1} - \mathbf{X}_i)^2 \right]. \quad (1.13)$$

The kinetic term alone would let the beads fly apart. The nearest-neighbor Hooke term is what makes them a string.

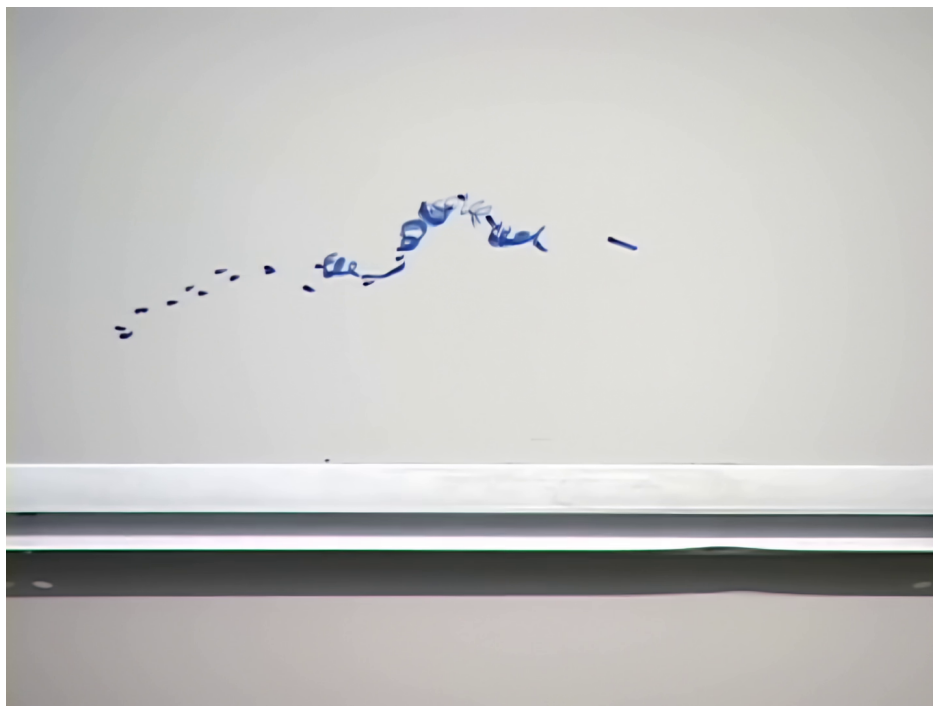


Figure 1.13: The first bead-and-spring sketch for an open string with two transverse coordinates.

Lecture frame: 01:00:30.

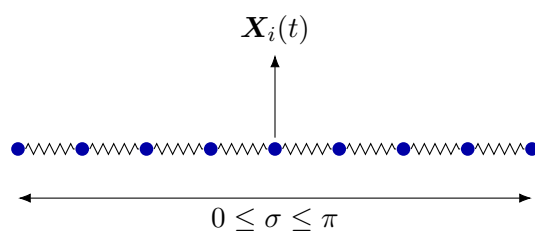
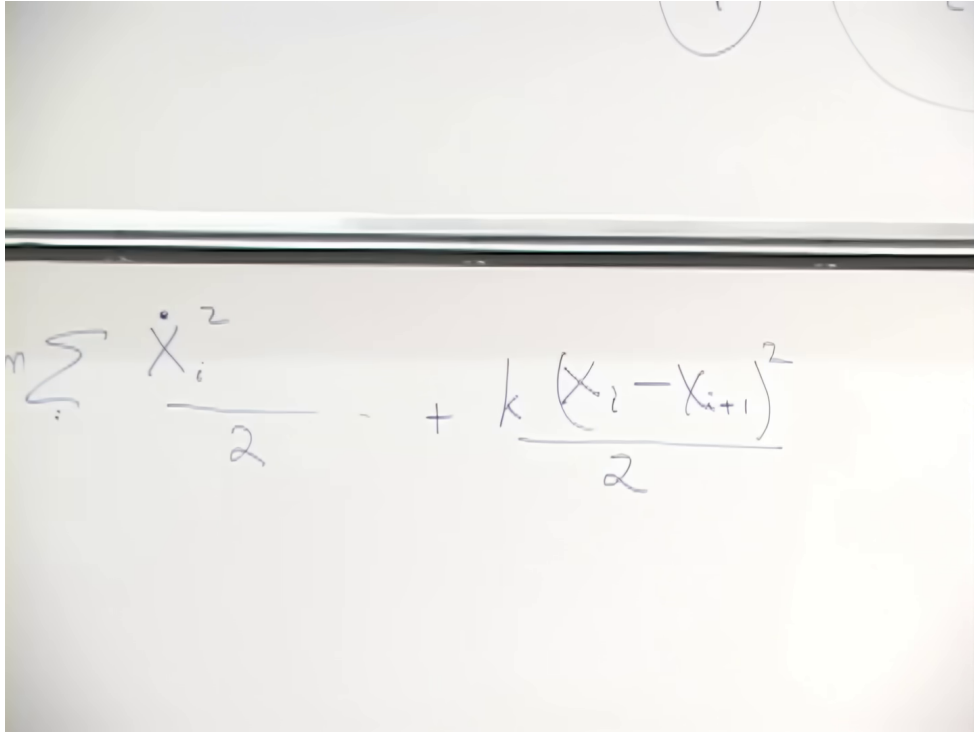


Figure 1.14: Discrete regulator for the open string. The bead mass and spring constant scale with the spacing before the continuum limit is taken.



$$n \sum_i \frac{\dot{X}_i^2}{2} + k \frac{(X_i - X_{i+1})^2}{2}$$

Figure 1.15: The discrete kinetic and Hooke terms written together before the continuum limit.

Lecture frame: 01:01:30.

As the number of beads increases, each bead mass μ_N must tend to zero so that the whole string retains finite inertia. At the same time the spring constant k_N must grow. A long strip of rubber is easier to stretch by a fixed amount than a tiny segment of the same material; neighboring points that are brought closer together must therefore be connected more stiffly if the continuum material is to remain the same.

Introduce a coordinate σ along the string. Its numerical range is a convention. Choosing

$$0 \leq \sigma \leq \pi \tag{1.14}$$

is convenient because a closed string can later be parameterized from 0 to 2π . The sum becomes an integral, and the nearest-neighbor difference becomes a derivative:

$$\sum_i \longrightarrow \int_0^\pi d\sigma, \quad \mathbf{X}_{i+1} - \mathbf{X}_i \longrightarrow \partial_\sigma \mathbf{X}. \tag{1.15}$$

After absorbing the spacing-dependent constants into the normalization of t , σ , and $\mathbf{X}$, the continuum Hamiltonian has the form

$$H_{\text{string}} = \frac{1}{2} \int_0^\pi d\sigma \left[(\partial_t \mathbf{X})^2 + (\partial_\sigma \mathbf{X})^2 \right]. \tag{1.16}$$

The image shows a whiteboard with handwritten mathematical expressions. At the top, there is a sum over beads: $= m \sum_i \dot{X}_i^2 + k \frac{(X_i - X_{i+1})^2}{2}$. An arrow points from this sum to a new expression below it, which is an integral from 0 to π over $d\sigma$ of a term in brackets: $\left[\frac{\dot{X}^2(\sigma)}{2} + \left(\frac{\partial X}{\partial \sigma} \right)^2 \right]$.

Figure 1.16: The board passage from the bead sum to $\int_0^\pi d\sigma$, with the nearest-neighbor difference replaced by $\partial_\sigma \mathbf{X}$.

Lecture frame: 01:05:00.

The first term is kinetic energy in the two transverse directions. The second measures stretching along the string. Both terms carry a plus sign because this is the Hamiltonian. In the Lagrangian the potential term would enter with a minus sign. The precise coefficients depend on conventions and were not fixed at this stage; the kinetic-plus-stretching structure is what will matter.

Separate the center-of-mass motion from the relative motion. Translating every point by the same amount moves the particle as a whole. Oscillations, bending, and relative stretching are internal motions. A proton, an atom, or any other composite object can be called a particle even though its constituents move internally. The string is no different in that respect.

1.8 Why the Hamiltonian measures mass squared

For a slowly moving object, rest energy is proportional to its mass. In the infinite-momentum description, however, the internal contribution appears as $M^2/(2P_z)$. Once the common longitudinal momentum and its associated time rescaling have been fixed, the internal string Hamiltonian is proportional to the invariant mass squared of the whole object:

$$H_{\text{int}} \propto M^2. \quad (1.17)$$

²Editorial clarification: More explicitly, light-cone kinematics gives an invariant relation of the form $M^2 =$

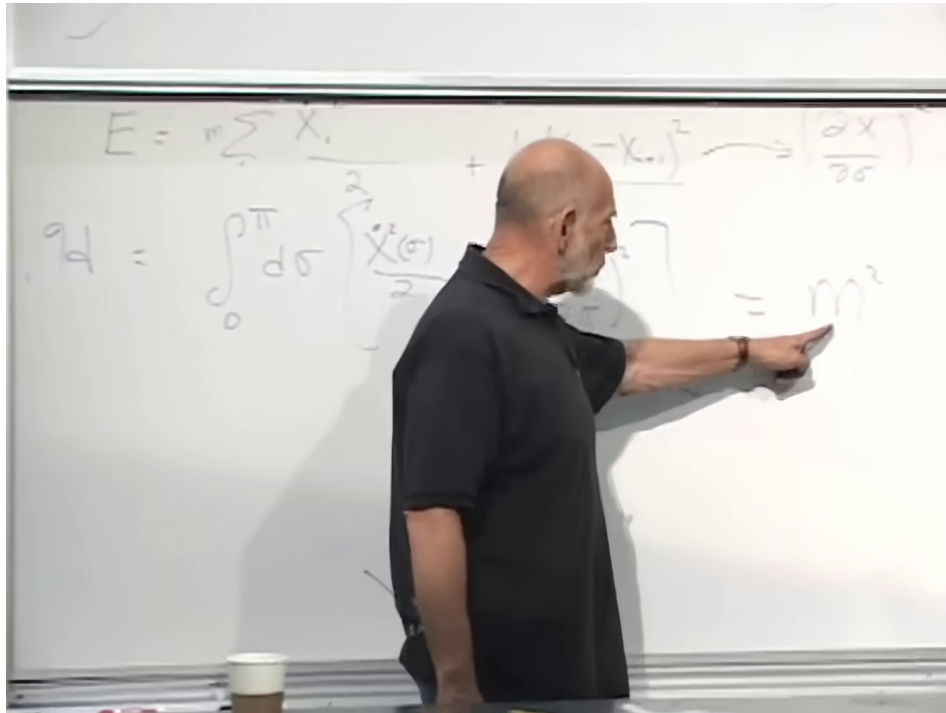


Figure 1.17: The board identifies the internal string energy with the mass squared of the complete string state.

Lecture frame: 01:10:30.

A vibrating string is a collection of coupled oscillators. Its complete quantization is postponed, but the important spectral fact is already visible. Harmonic-oscillator energies rise in discrete steps. Since those internal energies contribute to M^2 , string excitations naturally produce discrete steps in mass squared. This is the structure that the Regge plots announced at the beginning.

Question from the audience. Is M the mass of the whole string, and would doubling that mass quadruple the corresponding M^2 contribution? ^a

Response. Yes. M is the rest mass of the complete string state, including vibration, stretching, and every other internal contribution. The effective boosted Hamiltonian is proportional to M^2 , so a state with twice the mass has four times that mass-squared value.

^aLecture timestamp: 01:12:10.

Question from the audience. How can this description include a photon, whose rest mass is zero? ^a

Response. A massless state corresponds to $M = 0$. How the positive oscillator contributions can combine to produce such a state is a significant question for the quantized theory, not something settled by the classical model here. At this point M simply denotes the total rest mass, and quantum mechanics has not yet been introduced.

^aLecture timestamp: 01:12:52.

$2P^+P^- - P_\perp^2$. In the center-of-mass transverse frame and at fixed P^+ , the light-cone Hamiltonian P^- is therefore proportional to M^2 . The lecture suppresses these normalization details and keeps the proportionality needed for the argument.

Now stretch a nonmoving string uniformly to physical length L . The kinetic term vanishes, and over the parameter interval of length π ,

$$|\partial_\sigma \mathbf{X}| \simeq \frac{L}{\pi}. \quad (1.18)$$

Consequently,

$$\begin{aligned} H_{\text{stretch}} &\propto \int_0^\pi d\sigma (\partial_\sigma \mathbf{X})^2 \\ &\simeq \pi \left(\frac{L}{\pi} \right)^2 \propto L^2. \end{aligned} \quad (1.19)$$

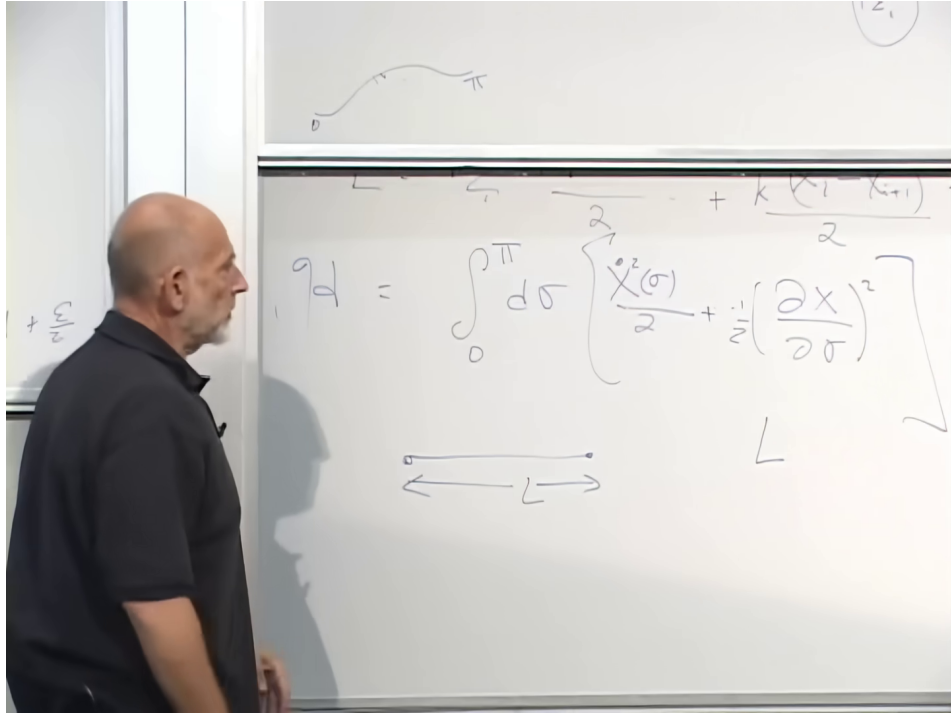


Figure 1.18: The uniformly stretched configuration used to infer $H_{\text{stretch}} \propto L^2$ and hence $M \propto L$.
 Lecture frame: 01:15:30.

This is Hooke-law behavior in the effective transverse mechanics. But the quantity on the left is proportional to M^2 , not M . It follows that

$$M^2 \propto L^2, \quad M \propto L. \quad (1.20)$$

In the rest frame the energy is $E_{\text{rest}} = M$, in units $c = 1$. Writing the constant of proportionality as the string tension T ,

$$E_{\text{rest}} = TL. \quad (1.21)$$

The result matches the flux-tube picture. A uniform electric or magnetic field carries a uniform energy density. If its flux is confined to a tube of fixed cross section, each additional unit of length adds the same amount of energy. Equivalently, a constituent picture in which new degrees of freedom fill each newly opened gap has fixed energy per unit length. String tension is that energy per unit length.

This conclusion is indirect. It does not claim that an arbitrary rubber band driven to relativistic speeds will automatically acquire $E = TL$. The argument begins with the transverse nonrelativistic description that follows from the large boost, identifies its internal energy with M^2 , and only then translates the result back to rest-frame energy.

1.9 The same object in two frames

The distinction between M and M^2 is easy to lose, so it is worth placing the two expansions side by side. With $c = 1$,

$$E = \sqrt{p^2 + M^2} \simeq M + \frac{p^2}{2M}, \quad p \ll M, \quad (1.22)$$

$$E = \sqrt{P_z^2 + M^2} \simeq P_z + \frac{M^2}{2P_z}, \quad P_z \gg M. \quad (1.23)$$

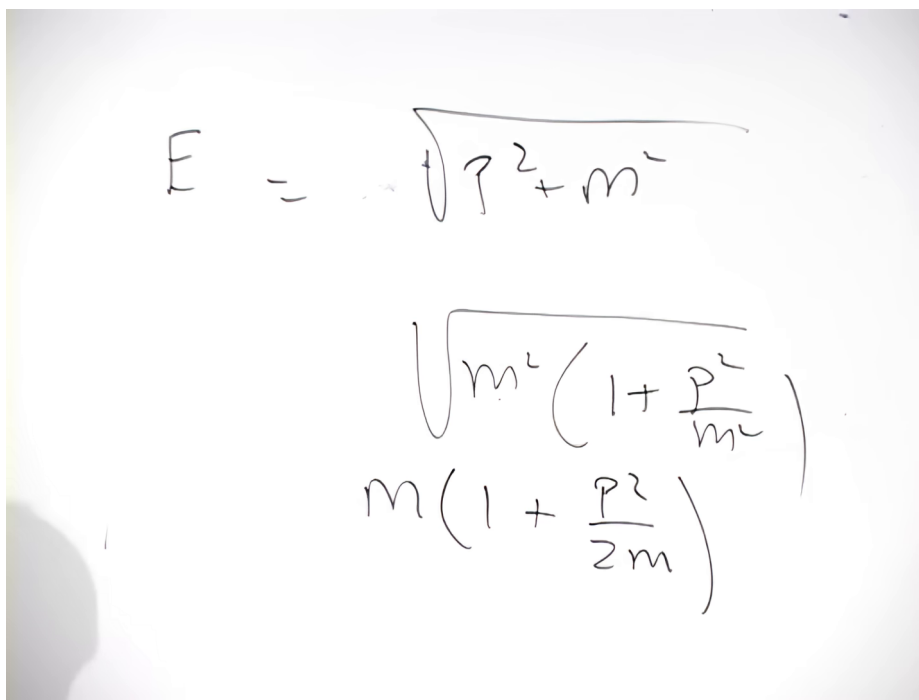


Figure 1.19: The small-momentum expansion repeated after the break, where internal rest energy is read as M .

$$\sqrt{P^2 \left(1 + \frac{m^2}{P_z}\right)}$$

$$P \approx \frac{M^2 + P_x^2 + P_y^2}{2P}$$

Figure 1.20: The complementary large-momentum expansion, where the residual energy is proportional to $M^2/(2P_z)$.

Lecture frame: 01:31:00.

The first expansion identifies rest energy with mass and ordinary inertia with M . In the second, the conserved P_z term is removed, transverse inertia is P_z , and internal energy is represented by $M^2/(2P_z)$. The mass remains proportional to the string length in either frame. What changes is which function of the mass appears as the residual Hamiltonian.

Question from the audience. A Lorentz boost should not change the physics. Why does the boosted string appear to obey a different energy law? ^a

Response. It does not obey different physics. In both descriptions $M \propto L$ and $M^2 \propto L^2$. In the rest frame the relevant energy is M ; after a large boost the conserved P_z is separated off and the remaining Hamiltonian is proportional to M^2 . The boost changes the useful description, not the invariant mass-length relation.

^aLecture timestamp: 01:25:35.

1.10 A superconducting model of confinement

A superconductor expels magnetic field from its bulk. If magnetic flux is trapped through a sufficiently narrow channel before the material becomes superconducting, the flux cannot spread sideways and remains confined to a tube. A useful Gedanken experiment is to drill a narrow hole through a normal-state sample, insert a long bar magnet through it, cool the sample, and then withdraw the magnet. The construction is schematic rather than a practical laboratory recipe. Its purpose is to isolate a flux line whose field strength and energy density are approximately uniform along its length.

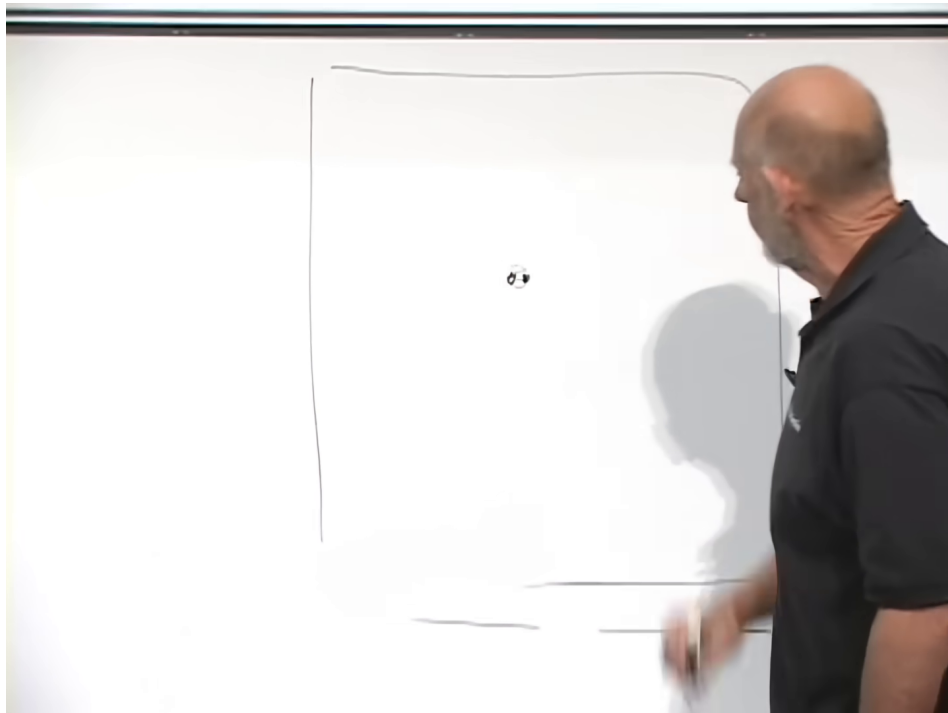


Figure 1.21: The superconducting flux-line construction and the corresponding monopole–antimonopole confinement picture.

Lecture frame: 01:35:30.

Suppose, as a further thought experiment, that a magnetic monopole and antimonopole could be placed in the superconducting medium and separated. The magnetic flux between them would be squeezed into a tube. Since each unit of tube length carries the same energy,

$$V(R) = T_{\text{flux}} R. \quad (1.24)$$

The force $F = -dV/dR$ has constant magnitude. Pulling the pair apart becomes increasingly expensive, so the pair is confined. This is the behavior sought for quarks in a hadron.

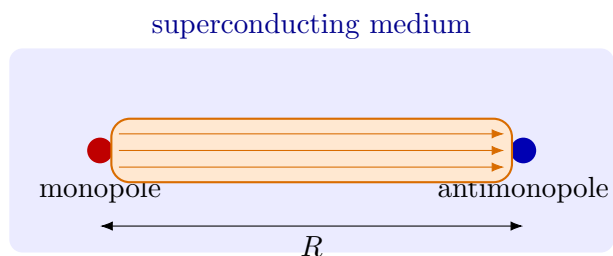


Figure 1.22: Flux confinement gives a uniform tube and a linear potential $V(R) = T_{\text{flux}} R$.

Question from the audience. What plays the role of the superconducting medium in the hadron?^a

Response. The proposed analogy reverses the electric and magnetic roles. An ordinary superconducting condensate is electrically charged and confines magnetic flux. In the dual-superconductor picture of QCD, a condensate of chromomagnetic objects confines the chromoelectric charge of quarks. Such objects do not appear as isolated particles because they are represented as condensed in the vacuum.

^aLecture timestamp: 01:36:57.

3

Question from the audience. Is energy proportional to length a general rule for any ordinary string made highly relativistic?^a

Response. No. One would first need a consistent relativistic theory of that material, and its behavior would depend on the details. Linear energy is special to a configuration whose density stays fixed as it is extended, or to flux prevented from spreading transversely. Pulling it apart creates more constituents or more uniform field-filled length rather than merely increasing the separation of a fixed set of molecules.

^aLecture timestamp: 01:38:10.

1.11 Transverse degrees of freedom and the size of a particle

The light-cone construction writes the string directly in terms of its transverse coordinates. There is no separate, freely specifiable longitudinal oscillator analogous to each transverse oscillator. This reduction by one direct spatial dimension is one of the remarkable features of the formalism; the complete theory nevertheless has to satisfy Lorentz invariance.

Question from the audience. What happens to residual motion along the z -direction? ^a

Response. In the light-cone description, longitudinal motion is not introduced as an independent degree of freedom. The physical oscillations are described transversely, while constraints determine the longitudinal information. This economy is related in spirit to the idea that gravitating systems may admit a description with one fewer direct dimension, although that connection is not derived here.

^aLecture timestamp: 01:40:10.

Question from the audience. How many points are actually known on the Regge plots? ^a

Response. The exact maximum depends on the family; a rough number is about seven. The sequence is not experimentally endless. The relevant fact is that the known points do not display the kind of visible failure or breakup endpoint familiar from ordinary macroscopic material.

^aLecture timestamp: 01:41:15.

Hadronic phenomenology supplied the historical motivation, but the mathematical string theory is not required to reproduce hadrons in every detail. Once the formalism is built, one asks what it

³Editorial clarification: The dual-superconductor description is an influential confinement mechanism and is presented here in the historical language of the lecture. It should be read as a physical picture of the QCD vacuum, not as a complete derivation of confinement from first principles.

actually describes. Its massless spin-two state and other properties point more naturally toward gravity.

The apparent contrast between hadrons and electrons can be quantified with the energy of a rigid rotation,

$$E_{\text{rot}} = \frac{J^2}{2I}, \quad I \sim MR^2. \quad (1.25)$$

Angular momentum is quantized. Writing J^2 schematically as $\hbar^2 n^2$, the first rotational step is of order

$$\Delta E_{0 \rightarrow 1} \sim \frac{\hbar^2}{2I}. \quad (1.26)$$

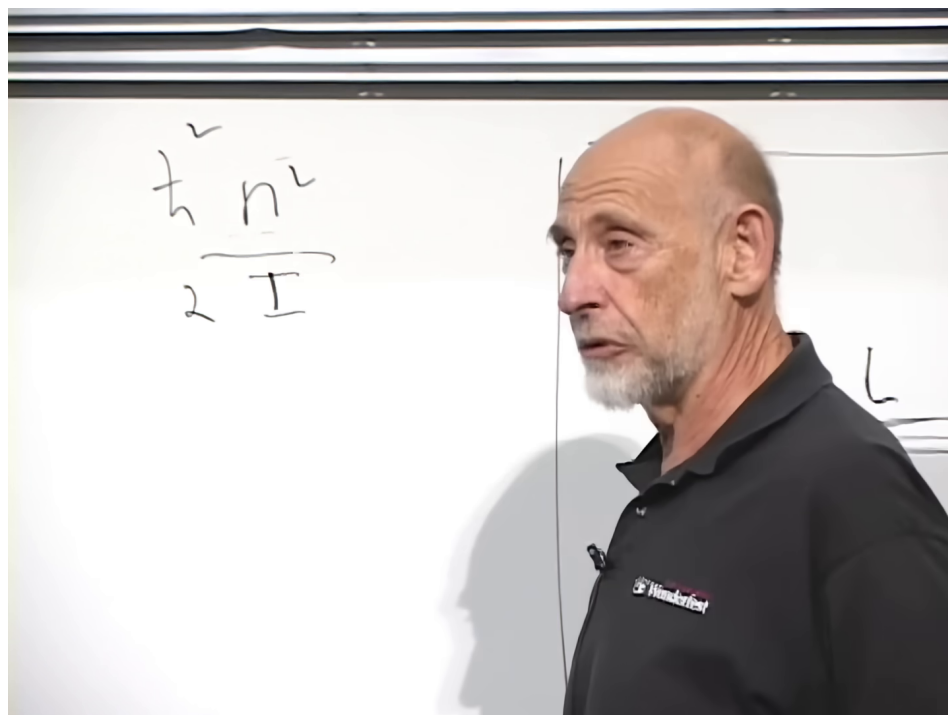


Figure 1.23: The moment-of-inertia estimate for the first rotational excitation, $\Delta E \sim \hbar^2/(2I)$.

Lecture frame: 01:44:00.

For fixed mass, a smaller radius means a much smaller moment of inertia and a much larger excitation energy. Hadrons are large on the scale of their internal dynamics, so a rotational excitation costs roughly a GeV and is accessible in particle collisions. If electrons possess any spatial substructure, experimental bounds make its scale far smaller, and rotational excitations would lie far beyond available energies. A basketball has the opposite extreme: its moment of inertia is so large that one quantum of angular momentum costs an imperceptibly small energy. Atomic rotational or orbital steps lie between these extremes, commonly at the scale of electron volts.

Question from the audience. Could an apparently elementary electron still have spatial structure and rotational excitations?^a

Response. It could only hide them at very high energy. The cost of one unit of angular momentum scales as $1/I$, and $I \sim MR^2$. An exceedingly small radius therefore pushes the first excitation far away on the mass-squared plot. The absence of observed electron excitations means either that the electron is pointlike or that any compositeness scale is much smaller than present experiments can resolve.

^aLecture timestamp: 01:42:04.

The opening Regge plot can now be read dynamically. Hadrons reveal accessible rotational and vibrational states because they are extended. The equal steps in their mass-squared spectrum are the signature expected when an oscillator Hamiltonian is identified with M^2 . Their approximate linear energy with length is the signature of a flux tube with constant tension. Spectroscopy, scattering duality, world-sheet geometry, and light-cone mechanics are four descriptions of the same organizing idea: an extended relativistic object whose natural internal variable is mass squared.

Chapter 2

From Boundary Conditions to the Open-String Spectrum

Overview. Our aim is to quantize an open string and decide whether its states deserve to be called particles. The route is elementary but exacting. We first prepare the calculus and Fourier analysis of functions on $0 \leq \sigma \leq \pi$. We then replace pointlikeness by a spectral criterion for particle behavior, review the infinite-momentum description in which the transverse dynamics become nonrelativistic, and build the string from a chain of light masses and stiff springs. The continuum string is a wave system with free endpoints. Its Neumann modes decouple into harmonic oscillators with frequencies $1, 2, 3, \dots$; quantization therefore produces a discrete, gapped spectrum. String interactions finally force us beyond open strings: an open string can join its own endpoints and become closed.

2.1 Calculus on the open-string interval

The calculation will repeatedly turn discrete sums into integrals and nearest-neighbor differences into derivatives. It is better to put those facts in place before they interrupt the physics.

Let $x(\sigma)$ be a function on

$$0 \leq \sigma \leq \pi. \quad (2.1)$$

The notation anticipates string theory: σ labels position along the string, while x will become one of its transverse coordinates. Divide the interval into N pieces,

$$\Delta\sigma = \frac{\pi}{N}, \quad \sigma_i = i \Delta\sigma, \quad x_i = x(\sigma_i). \quad (2.2)$$

For a smooth limiting function,

$$\Delta x_i \equiv x_i - x_{i-1} = \frac{\partial x}{\partial \sigma} \Delta\sigma + O((\Delta\sigma)^2), \quad (2.3)$$

and the Riemann sum becomes

$$\sum_i x_i \Delta\sigma \longrightarrow \int_0^\pi x(\sigma) d\sigma. \quad (2.4)$$

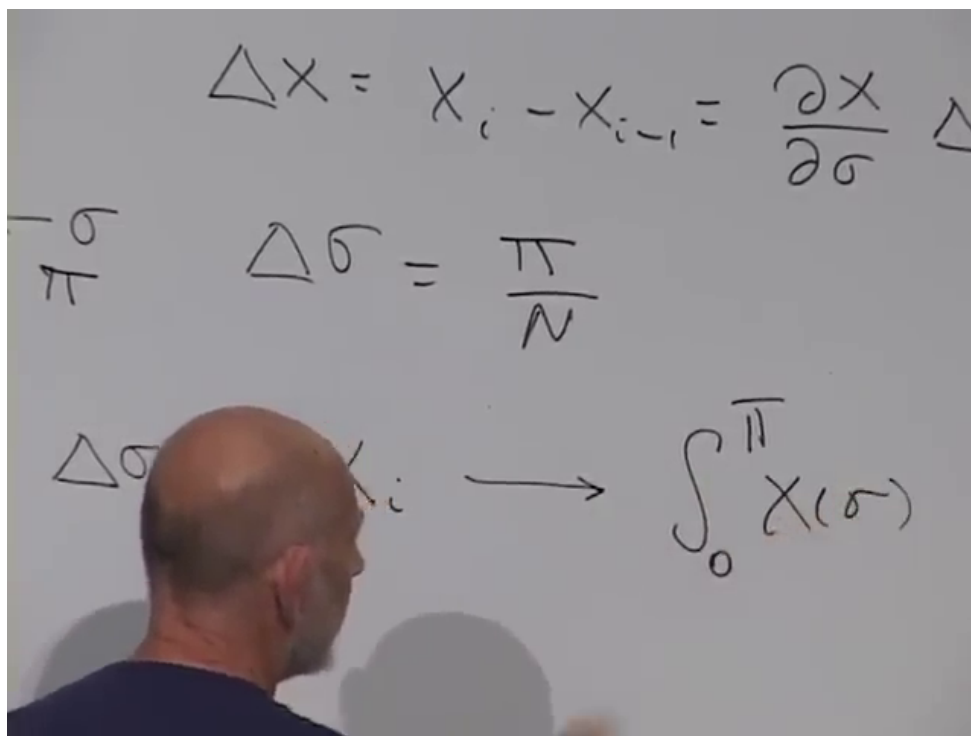


Figure 2.1: The finite-difference and Riemann-sum rules kept on the board for the later continuum limit.

Lecture frame: 00:04:20.

These are definitions familiar from calculus. Their importance here is in the scaling: Δx_i brings one power of $\Delta\sigma$, whereas a sum brings one inverse power when it becomes an integral.

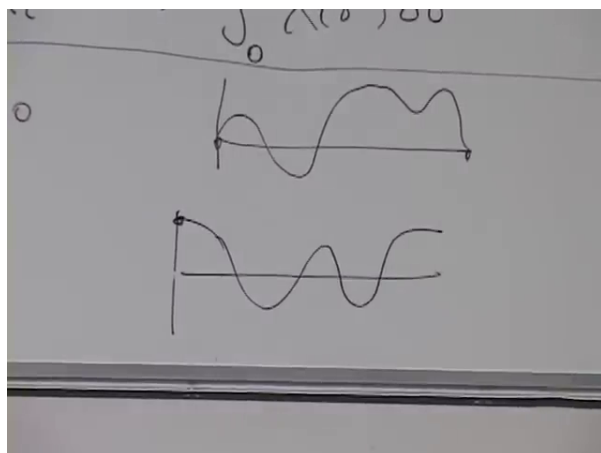
2.1.1 Dirichlet and Neumann endpoints

Fourier analysis on a finite interval begins with a choice at the endpoints. Two choices will recur.

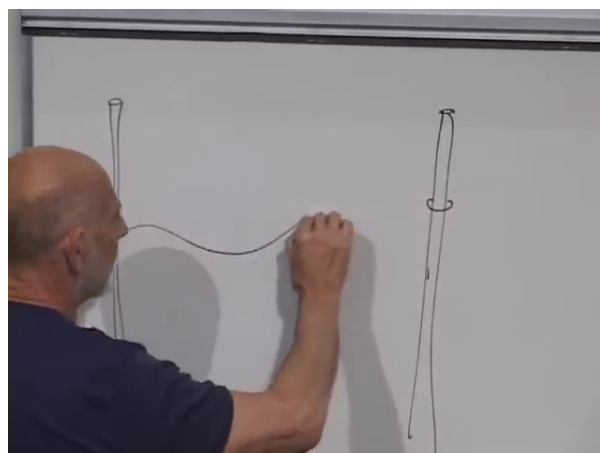
$$\text{Dirichlet:} \quad x(0) = x(\pi) = 0, \quad (2.5)$$

$$\text{Neumann:} \quad \partial_\sigma x|_{\sigma=0} = \partial_\sigma x|_{\sigma=\pi} = 0. \quad (2.6)$$

A violin string held at both ends gives the Dirichlet picture: its displacement vanishes at the boundaries. A Neumann function need not vanish there; it meets each boundary with zero slope. The two characteristic shapes are visible together on the board.



(a) Pinned and flat endpoint profiles.



(b) Light rings sliding on vertical poles.

Figure 2.2: Boundary conditions as shapes and as mechanics. The sliding-ring construction gives a free endpoint an unambiguous transverse motion.

Lecture frame: 00:08:50 and 00:12:15.

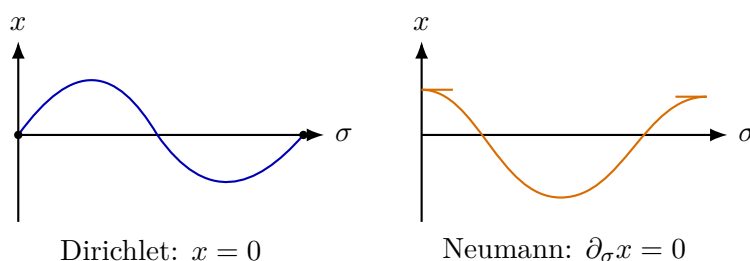


Figure 2.3: The value is fixed for Dirichlet data; the slope is fixed for Neumann data.

Question from the audience. Does an open or closed organ pipe supply a clean example of one of these boundary conditions?^a

Response. It can, but one must first say which acoustic variable is being expanded. Pressure perturbation and longitudinal displacement obey complementary idealized conditions at an open end. Rather than settle that side question at the board, use the sliding-ring model: nearly massless rings can move vertically on the poles, so the string endpoints are free and the Neumann picture is direct.

^aLecture timestamp: 00:09:19.

Question from the audience. Can the graph rise vertically and contain a sharp jump?^a

Response. Fourier analysis permits isolated jumps in a piecewise continuous function. A physical string is more restrictive. A jump disconnects neighboring points and makes the spatial derivative $\partial_\sigma x$ singular in the continuum energy derived below. We therefore restrict attention to profiles that are continuous and differentiable.

^aLecture timestamp: 00:09:53.

Question from the audience. Why should the open end of a pipe, or any free end, satisfy a boundary condition at all?^a

Response. The acoustic example depends on the chosen field and was left unresolved in the discussion. The mechanical example does not. A very light ring is free to slide but is still attached to the final string segment. We shall derive its condition from Newton's law rather than assume it: finite endpoint acceleration forces the slope to vanish.

^aLecture timestamp: 00:11:14.

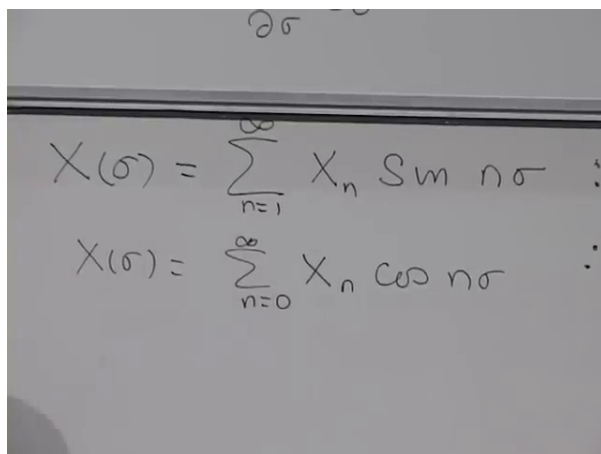
2.1.2 The Fourier bases

Once the endpoint class is fixed, the corresponding Fourier basis is fixed. On $0 \leq \sigma \leq \pi$,

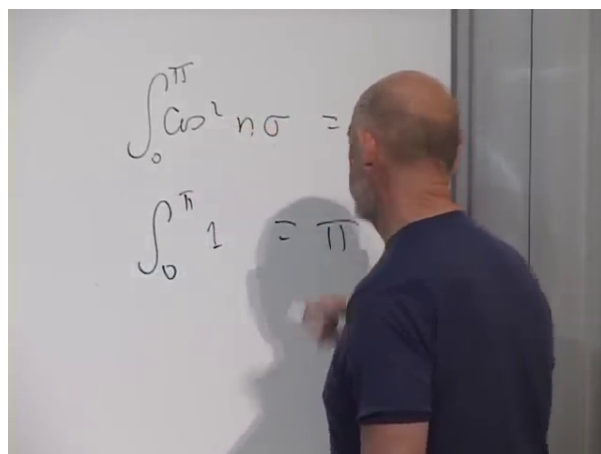
$$x_D(\sigma) = \sum_{n=1}^{\infty} x_n \sin(n\sigma), \quad (2.7)$$

$$x_N(\sigma) = \sum_{n=0}^{\infty} x_n \cos(n\sigma). \quad (2.8)$$

Every sine in (2.7) vanishes at both endpoints. Every cosine in (2.8) has zero endpoint derivative. The lower limits differ for a reason: $\sin(0\sigma) = 0$ contributes nothing, while $\cos(0\sigma) = 1$ is the constant, or zero, mode. It records the average value of a Neumann function.



(a) Sine and cosine expansions.



(b) Orthogonality and the exceptional constant mode.

Figure 2.4: The Fourier facts used later to diagonalize the string.

Lecture frame: 00:16:15 and 00:20:30.

Differentiation exchanges the two classes:

$$\begin{aligned} \partial_\sigma \sin(n\sigma) &= n \cos(n\sigma), \\ \partial_\sigma \cos(n\sigma) &= -n \sin(n\sigma). \end{aligned} \quad (2.9)$$

Thus the derivative of a Dirichlet series has Neumann behavior and the derivative of a Neumann series has Dirichlet behavior. Orthogonality then removes cross-terms between distinct modes:

$$\int_0^\pi \cos(n\sigma) \cos(m\sigma) d\sigma = \frac{\pi}{2} \delta_{nm}, \quad n, m \geq 1, \quad (2.10)$$

$$\int_0^\pi \sin(n\sigma) \sin(m\sigma) d\sigma = \frac{\pi}{2} \delta_{nm}, \quad n, m \geq 1, \quad (2.11)$$

$$\int_0^\pi 1 d\sigma = \pi. \quad (2.12)$$

The last line is the exceptional $n = m = 0$ cosine norm. A constant is not normalized like a nonzero harmonic, and that factor of two will distinguish the string's center-of-mass mode from its internal modes.

2.1.3 One oscillator formula to remember

After rescaling a coordinate so that its mechanical mass is one, a Hooke-law oscillator has

$$E_{\text{osc}} = \frac{1}{2} \dot{q}^2 + \frac{1}{2} \omega^2 q^2, \quad (2.13)$$

$$L_{\text{osc}} = \frac{1}{2} \dot{q}^2 - \frac{1}{2} \omega^2 q^2. \quad (2.14)$$

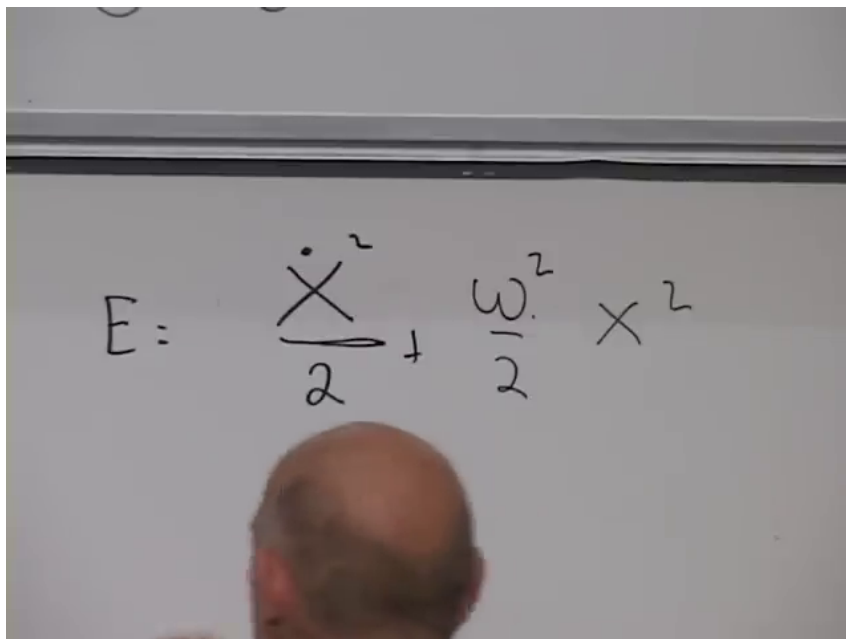


Figure 2.5: The oscillator energy, written so that the coefficient of q^2 immediately identifies ω^2 .

Lecture frame: 00:24:00.

Classically, $\omega = \sqrt{\kappa/m}$; the coordinate rescaling has set $m = 1$. Quantum mechanically,

$$E_Q = \hbar\omega \left(Q + \frac{1}{2} \right), \quad Q = 0, 1, 2, \dots \quad (2.15)$$

For level spacings we may subtract the common ground-state term and say that each quantum costs $\hbar\omega$. The string calculation will initially do exactly that and postpone the total zero-point energy.

Question from the audience. How is the oscillator frequency read from the energy formula?^a

Response. Before rescaling it is $\omega = \sqrt{\kappa/m}$. With unit mass the potential term is $\frac{1}{2}\omega^2 q^2$, so its quadratic coefficient is the square of the frequency.

^aLecture timestamp: 00:23:21.

2.2 What makes a composite system a particle?

A string is extended and will be modeled by many coupled pieces. In what sense can one of its quantum states nevertheless be a particle? Pointlikeness cannot be the decisive criterion. Protons and neutrons are composite. Electrons and photons are accompanied by quantum-field dressing even though no intrinsic substructure has been resolved. A classical geometric point surrounded by nothing is not the right image of a quantum excitation.

Nor do mass and center-of-mass position suffice. A sealed box of gas has both. A cup of coffee has both. We do not normally call either one a particle. Indivisibility also fails because a proton is divisible, and spin does not distinguish microscopic particles from macroscopic bodies.

The useful criterion is spectral isolation. Put the rest energy on a vertical axis and ask how much energy is required to reach another internal state. An electron has no experimentally accessible tower of electron excitations just above it. A proton does have resonances, but the first ones cost a few hundred MeV, a substantial fraction of the proton mass. Those states are separated enough to be individually identifiable.

A cup of coffee is the opposite. One can add an arbitrarily small amount of energy to a long-wavelength sound mode, nudge a molecule, or produce a phonon. There are enormously many states packed into a tiny interval. On laboratory scales they form a continuum. This is the practical distinction between a particle and *mush*.

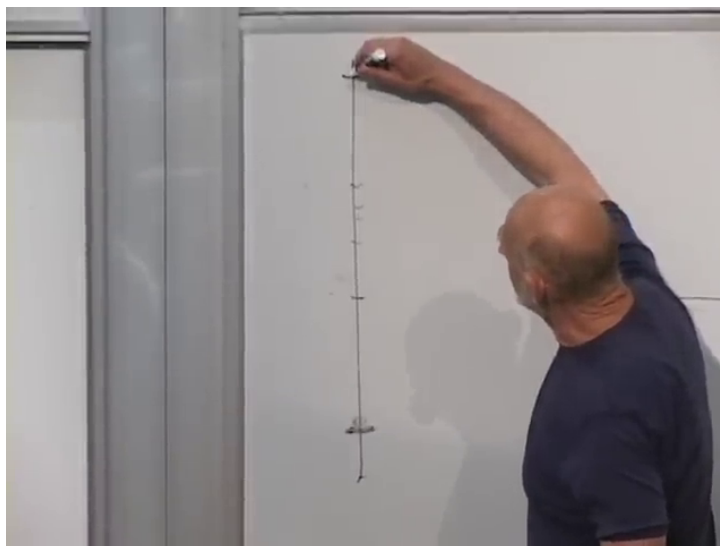


Figure 2.6: The board's vertical energy axis during the comparison of an isolated particle level with the dense excitations of ordinary matter.

Lecture frame: 00:31:10.

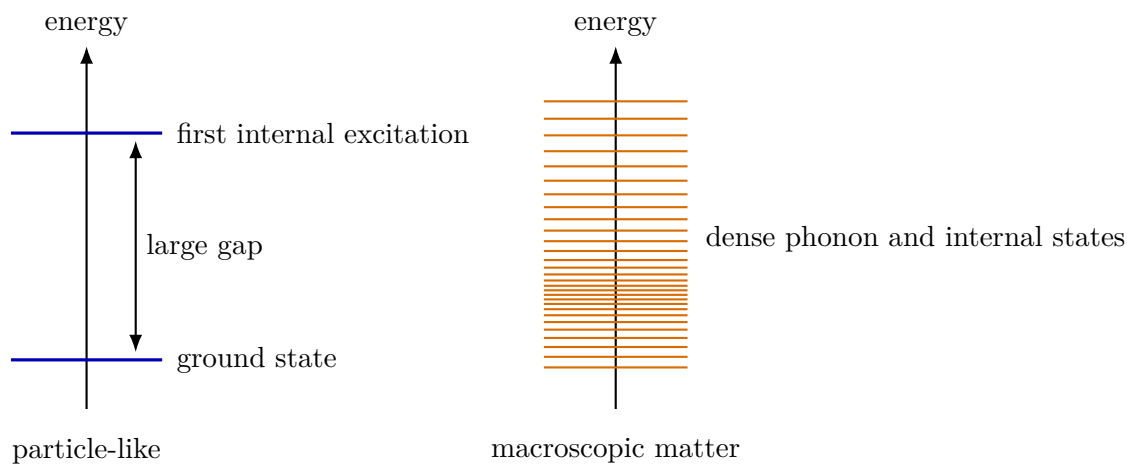


Figure 2.7: Spectral isolation, not literal pointlikeness, is the operational particle criterion used in the lecture.

This criterion turns the philosophical question into a calculation. Quantize the string, find its internal energy levels, and inspect their spacing. If the first excited string state lies near the Planck scale while the observed ground state is light, the gap exceeds present accelerator energies by more than twenty orders of magnitude. The string then behaves as an elementary particle throughout the accessible regime.

Question from the audience. Is particle behavior merely an issue of observational scale? If we zoom in far enough, do the levels cease to be separate?^a

Response. Zooming in may reveal that an object has many constituents, but it does not erase the spectrum. The relevant comparison is between the energy resolution of the experiment and the gap to the first internal excitation. A composite object can remain particle-like when that gap is enormous.

^aLecture timestamp: 00:33:33.

Question from the audience. Why is an ordinary violin string not a particle?^a

Response. One vibrational quantum carries an immeasurably small energy compared with the rest energy of the macroscopic string. Many quantum states therefore fit into a tiny relative interval. It belongs on the dense side of the spectral comparison, not the isolated side.

^aLecture timestamp: 00:34:37.

Question from the audience. Are photons and electrons point particles if quantum fields surround them with virtual fluctuations?^a

Response. The classical word *point* is doing too much work. The physical electron and photon are dressed field excitations, yet experiments have found no intrinsic constituent structure or finite form factor for either. What matters for the present argument is that neither has a dense tower of almost-degenerate internal states.^b

^aLecture timestamp: 00:35:40.

^bEditorial clarification: The lecture associates the electron's anomalous magnetic moment with a measurable quantum fuzz. In modern terminology the anomalous moment is a radiative dressing effect, not evidence for a resolved finite electron radius. Pointlike conventionally means that no intrinsic substructure or form factor has been detected down to the tested scale.

Question from the audience. Does the size of a photon depend on its frequency?^a

Response. Its wavelength certainly does. That is distinct from an intrinsic constituent size. Changing the frequency changes the wave packet and wavelength, not a resolved internal radius of the photon.

^aLecture timestamp: 00:36:49.

2.3 Exact transverse mechanics at large momentum

We now return to the relativistic difficulty. A rapidly rotating baton cannot be treated nonrelativistically merely by putting its center of mass at rest. Its ends still move relativistically relative to one another. The nonrelativistic approximation fails because of internal relative motion.

There is, however, a different limit. Choose a z -axis and boost the entire system to a very large positive longitudinal momentum P_z . Keep its transverse momentum

$$p_{\perp}^2 = p_x^2 + p_y^2 \tag{2.16}$$

finite. The transverse projection then admits an exactly nonrelativistic Hamiltonian form in the $P_z \rightarrow \infty$ limit, even if the unboosted internal motion was relativistic.

Question from the audience. Does infinite momentum mean that the massive system actually reaches the speed of light?^a

Response. No finite boost reaches it. One constructs quantities with well-defined limits as P_z becomes arbitrarily large. The phrase names a limiting procedure, not a finite physical state with $v = c$.

^aLecture timestamp: 00:40:43.

Question from the audience. Which axis must be chosen, and what if the object rotates about or across that axis?^a

Response. Any axis may be chosen and the whole system boosted along it. What remains as explicit mechanics is the projection onto the perpendicular plane. A rotating rod may look foreshortened or sweep through that plane, but its transverse projection has the nonrelativistic form.

^aLecture timestamp: 00:41:14.

Strings have an additional property that makes this representation unusually powerful. In light-front gauge their longitudinal internal coordinates are not independent oscillators. The constraints determine them from the transverse motion. Knowing the transverse string is therefore enough to reconstruct the physical degrees of freedom.

Question from the audience. Does the absence of an independent longitudinal coordinate mean that the string is merely trapped in a thin sandwich?^a

Response. Lorentz contraction does make its longitudinal extent small in the boosted frame, but the statement is stronger. Any residual longitudinal motion is constrained by the transverse motion; it cannot be specified independently.

^aLecture timestamp: 00:42:26.

Question from the audience. Is this reduction to transverse data related to the holographic principle?^a

Response. It is closely connected to the line of thought that led to holographic ideas, although pursuing that connection would take us away from the present calculation. Here it serves only to justify working with the transverse string.

^aLecture timestamp: 00:43:12.

2.3.1 Subtracting the boost and rescaling time

Let M be the invariant mass of the complete system, including constituent masses, kinetic energy, and binding energy. In units $c = 1$,

$$E = \sqrt{P_z^2 + p_\perp^2 + M^2}. \quad (2.17)$$

For P_z much larger than $p_\perp$ and M ,

$$E = P_z + \frac{p_\perp^2 + M^2}{2P_z} + O(P_z^{-3}). \quad (2.18)$$

$$E = P_z + \frac{p_{\perp}^2 + M^2}{2P_z}$$

Figure 2.8: The completed large- P_z expansion, with transverse momentum and invariant mass retained.

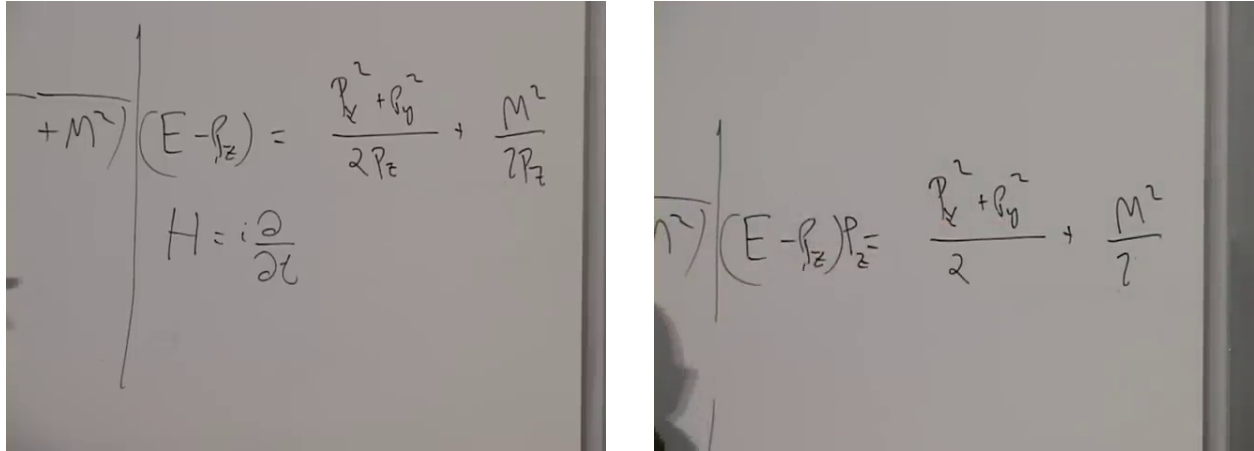
Lecture frame: 00:47:30.

The leading term is large, but it carries no information about the internal state. In a process that conserves P_z , such as a boosted atom changing between internal levels without changing its total longitudinal momentum, that term cancels from every energy difference. We therefore study

$$E - P_z = \frac{p_{\perp}^2 + M^2}{2P_z} + O(P_z^{-3}). \quad (2.19)$$

The denominator has a physical meaning. Since the Hamiltonian generates time evolution, the small residual energy says that internal evolution appears slow in laboratory time. This is time dilation. To retain the internal motion, introduce a rescaled evolution parameter τ , or equivalently multiply the residual Hamiltonian by P_z :

$$H_{\perp} \equiv \lim_{P_z \rightarrow \infty} P_z(E - P_z) = \frac{1}{2}p_{\perp}^2 + \frac{1}{2}M^2. \quad (2.20)$$


 (a) $H = i\partial_t$ and the slowed internal evolution.

(b) The finite transverse Hamiltonian after time rescaling.

Figure 2.9: From the time-dilated residual energy to a finite two-dimensional Hamiltonian.

Lecture frame: 00:49:20 and 00:52:40.

This is precisely the form of a unit-mass particle moving in two dimensions: kinetic energy $p_{\perp}^2/2$ plus an internal energy $M^2/2$. The unusual feature is the square. In this description, the quantity playing the role of rest or binding energy is proportional to invariant mass squared.¹

Question from the audience. What happened to the longitudinal momenta of the individual parts of a composite system?^a

Response. For a generic composite object one must retain the fractions of total P_z carried by its constituents. The formula above uses the total center-of-mass momentum. String theory is special because its constraints determine the relative longitudinal motion from the transverse degrees of freedom, so no independent longitudinal oscillator has to be added.

^aLecture timestamp: 00:51:30.

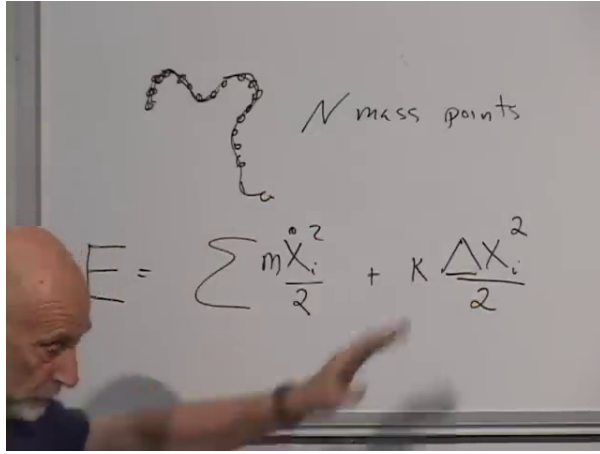
2.4 A discrete open string

The boosted string now moves explicitly only in the transverse plane. Label its two coordinates X and Y , and first approximate it by N mass points joined by $N - 1$ springs. If μ is the analog mass of each point and k the nearest-neighbor spring constant, the energy is

$$E_N = \sum_i \left[\frac{\mu}{2} (\dot{X}_i^2 + \dot{Y}_i^2) + \frac{k}{2} ((\Delta X_i)^2 + (\Delta Y_i)^2) \right]. \quad (2.21)$$

Every X -term has an identical Y -term because the transverse plane has two directions. The index i labels a point along the chain; it is not a spatial coordinate.

¹Editorial clarification: With conventional light-front variables $P^{\pm} = (E \pm P_z)/\sqrt{2}$, the exact mass-shell relation is $2P^+P^- = p_{\perp}^2 + M^2$. Equation (2.20) is the lecture's large- P_z normalization of the same invariant relation.



(a) The bead chain and its discrete energy.

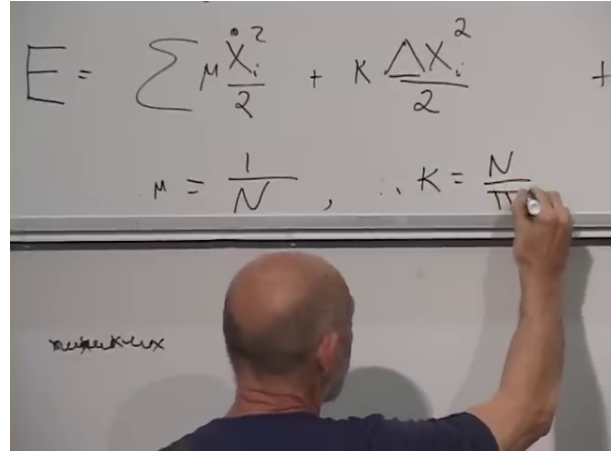

 (b) The large- N scaling of μ and k .

Figure 2.10: The regulated open string before taking the continuum limit.

Lecture frame: 00:57:42 and 01:02:40.

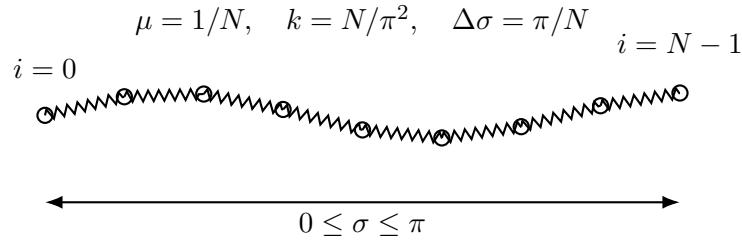


Figure 2.11: A clean regulator picture. The points become lighter while the microscopic springs become stiffer as the chain is refined.

The parameter σ labels position along the chain. Choosing $0 \leq \sigma \leq \pi$ is a convention, not a physical length measurement. It reminds us that the string has two endpoints. Later a closed string will be labeled periodically over $0 \leq \sigma < 2\pi$.

The continuum limit must keep the collective mechanics finite. Fixed bead mass would make the total analog mass diverge, so choose

$$\mu = \frac{1}{N}. \quad (2.22)$$

Then $N\mu = 1$. This is not the relativistic rest mass M ; it is the inertial coefficient in the transverse analog model.

Springs in series become softer as their number increases. To preserve a finite stiffness for the whole chain, each microscopic spring must become stiffer:

$$k = \frac{N}{\pi^2} = \frac{1}{\pi \Delta\sigma}. \quad (2.23)$$

Only $k \propto N$ is essential. The factor of π^2 is a normalization chosen to remove stray factors of π from the continuum formulas.

Question from the audience. Is μ the physical rest mass of each string point?^a

Response. No. It is the nonrelativistic analog mass appearing in the transverse chain. Calling it M would confuse it with the invariant rest mass of the entire relativistic string state. Its only present role is to produce the correct continuum inertia.

^aLecture timestamp: 01:10:34.

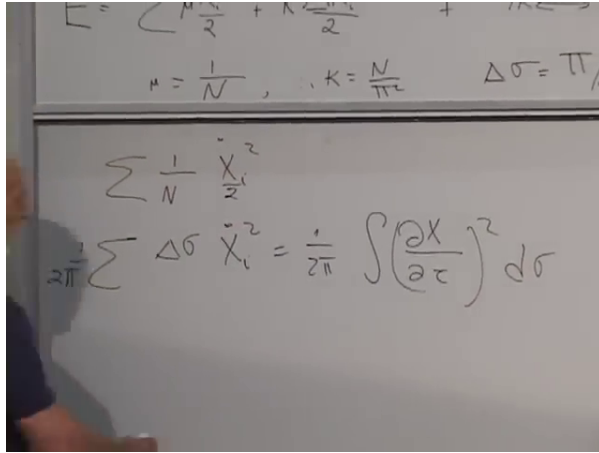
2.5 The continuum wave system

Apply the rules prepared at the beginning. Since $1/N = \Delta\sigma/\pi$,

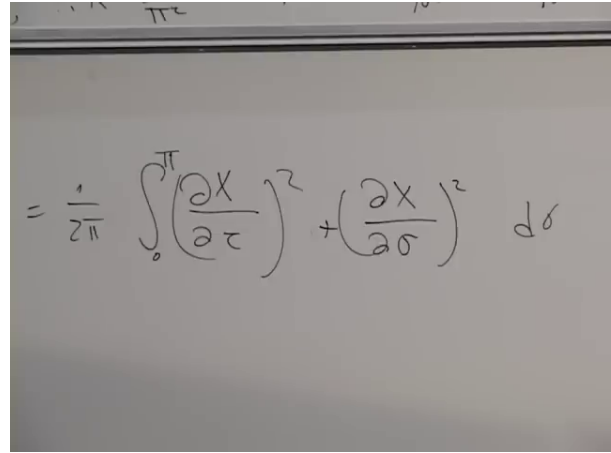
$$\sum_i \frac{\mu}{2} \dot{X}_i^2 = \frac{1}{2\pi} \sum_i \dot{X}_i^2 \Delta\sigma \longrightarrow \frac{1}{2\pi} \int_0^\pi (\partial_\tau X)^2 d\sigma, \quad (2.24)$$

$$\sum_i \frac{k}{2} (\Delta X_i)^2 \longrightarrow \frac{1}{2\pi} \int_0^\pi (\partial_\sigma X)^2 d\sigma. \quad (2.25)$$

For the second line, use $\Delta X_i = \partial_\sigma X \Delta\sigma$ together with $k = 1/(\pi\Delta\sigma)$. The same calculation applies to Y .



(a) Sum-to-integral and difference-to-derivative replacements.



(b) The completed continuum energy.

Figure 2.12: The discrete chain becoming a field on the σ -interval.

Lecture frame: 01:05:40 and 01:07:00.

The resulting energy and Lagrangian are

$$E = \frac{1}{2\pi} \int_0^\pi d\sigma \left[(\partial_\tau X)^2 + (\partial_\tau Y)^2 + (\partial_\sigma X)^2 + (\partial_\sigma Y)^2 \right], \quad (2.26)$$

$$L = \frac{1}{2\pi} \int_0^\pi d\sigma \left[(\partial_\tau X)^2 + (\partial_\tau Y)^2 - (\partial_\sigma X)^2 - (\partial_\sigma Y)^2 \right]. \quad (2.27)$$

Figure 2.13: The continuum Lagrangian with kinetic and stretching terms carrying opposite signs.
Lecture frame: 01:08:48.

If we rename σ as an ordinary spatial coordinate and X as a field ϕ , these are the energy and Lagrangian of a one-dimensional wave field. Varying (2.27) gives

$$\left(\partial_\tau^2 - \partial_\sigma^2 \right) X = 0, \quad \left(\partial_\tau^2 - \partial_\sigma^2 \right) Y = 0. \quad (2.28)$$

Nothing in the immediate argument requires us to solve these equations directly. The useful picture is simply that waves travel left and right along the string and reflect at the endpoints. A Dirichlet reflection reverses the sign of the displacement; a Neumann reflection preserves it.

Question from the audience. What exactly is the time variable τ ?^a

Response. It is the rescaled time used to follow the internal motion in the large- P_z description. A clock boosted with the system runs slowly relative to laboratory time; τ removes that dilation so the internal oscillations retain a finite rate. For this transverse mechanics it can be treated simply as time.^b

^aLecture timestamp: 01:08:54.

^bEditorial clarification: The lecture informally calls τ proper time. For an extended string, the precise light-front evolution parameter is a gauge-fixed rescaled time rather than a single invariant proper time shared by every point. The intended time-dilation argument is unchanged.

When the center of mass is at rest in the transverse plane, the internal part of (2.26) is, in the normalization of (2.20), proportional to M^2 . That is why quantizing this apparently ordinary wave system will reveal the relativistic mass spectrum.

Question from the audience. Does $P_z(E - P_z)$ have a name?^a

Response. It is commonly called a light-cone or light-front energy. Light front better describes the null plane used as the evolution surface; the traditional terminology is not important for the calculation.

^aLecture timestamp: 01:11:00.

Question from the audience. Is the word cone referring to the usual spacetime light cone, and did Dirac introduce this terminology?^a

Response. The null-coordinate construction can be drawn with 45-degree axes, but the evolution surface here is a light front rather than the full expanding cone. Dirac developed related forms of relativistic dynamics; the historical naming question was left open and does not affect the equations.

^aLecture timestamp: 01:11:17.

2.6 Why a free endpoint is Neumann

We can now derive rather than postulate the endpoint condition. Newton's law is legitimate because the boosted transverse mechanics has exactly the nonrelativistic form. An interior mass point is pulled by springs on both sides. The final point is pulled only by the spring joining it to its neighbor.

Question from the audience. Why are we allowed to use Newton's law while discussing a relativistic string?^a

Response. We are not applying Newtonian mechanics to the full unboosted spacetime motion. We are applying it to the transverse variables in the infinite-momentum description, whose Hamiltonian has exactly the nonrelativistic form.

^aLecture timestamp: 01:14:52.

At the right endpoint, Hooke's law gives, up to orientation,

$$F_{\text{end}} = -k \Delta X = -\frac{1}{\pi \Delta \sigma} (\partial_\sigma X \Delta \sigma) = -\frac{1}{\pi} \partial_\sigma X. \quad (2.29)$$

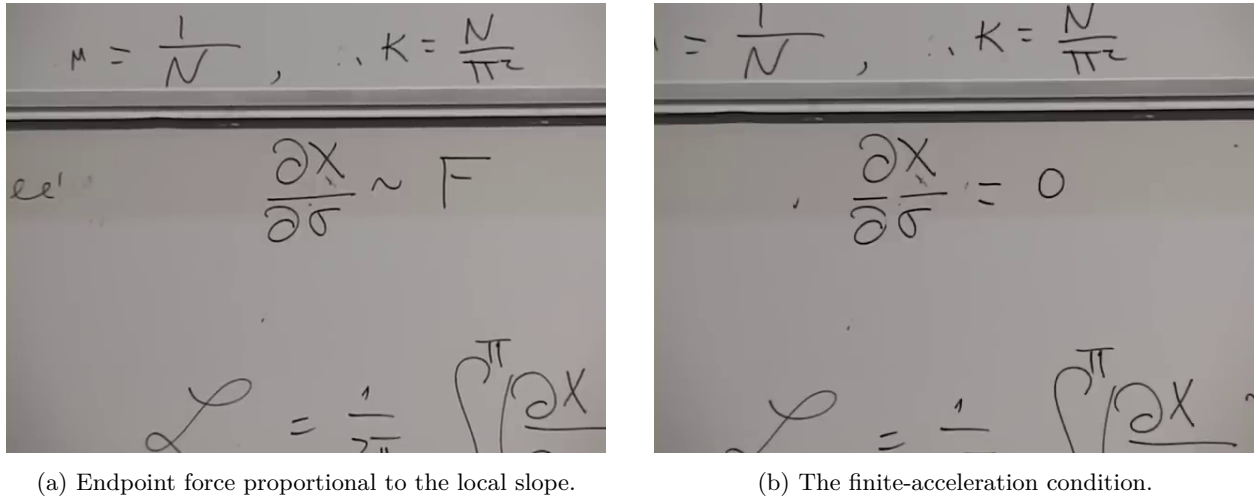


Figure 2.14: The Newton-law derivation of the free-end boundary condition.

Lecture frame: 01:18:20 and 01:20:10.

Newton's equation for the final point is

$$\mu \partial_\tau^2 X = F_{\text{end}}, \quad \mu = \frac{1}{N} = \frac{\Delta\sigma}{\pi}. \quad (2.30)$$

Combining (2.29) and (2.30),

$$\partial_\sigma X = -\Delta\sigma \partial_\tau^2 X \quad \text{at the right endpoint,} \quad (2.31)$$

with the force sign reversed at the left endpoint. If $\partial_\sigma X$ approached a nonzero constant while $\Delta\sigma \rightarrow 0$, the endpoint acceleration would diverge like N . A sensible continuum limit with finite acceleration therefore requires

$$\partial_\sigma X|_{0,\pi} = 0, \quad \partial_\sigma Y|_{0,\pi} = 0. \quad (2.32)$$

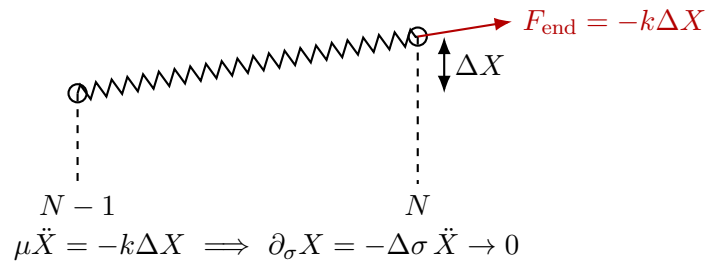


Figure 2.15: The last point has vanishing mass in the regulator limit. A finite acceleration is possible only when the endpoint slope vanishes.

The same argument at both ends and in both transverse directions produces Neumann boundary conditions for a free open string.

2.7 Normal modes of the open string

We now have a complete classical system: a Lagrangian and its boundary conditions. Quantization becomes simple after Fourier analysis. Because the endpoints are Neumann,

$$X(\sigma, \tau) = \sum_{n=0}^{\infty} X_n(\tau) \cos(n\sigma), \quad (2.33)$$

$$Y(\sigma, \tau) = \sum_{n=0}^{\infty} Y_n(\tau) \cos(n\sigma). \quad (2.34)$$

The cosines contain all σ -dependence. The coefficients contain all time-dependence; the string wiggles because $X_n(\tau)$ and $Y_n(\tau)$ change.

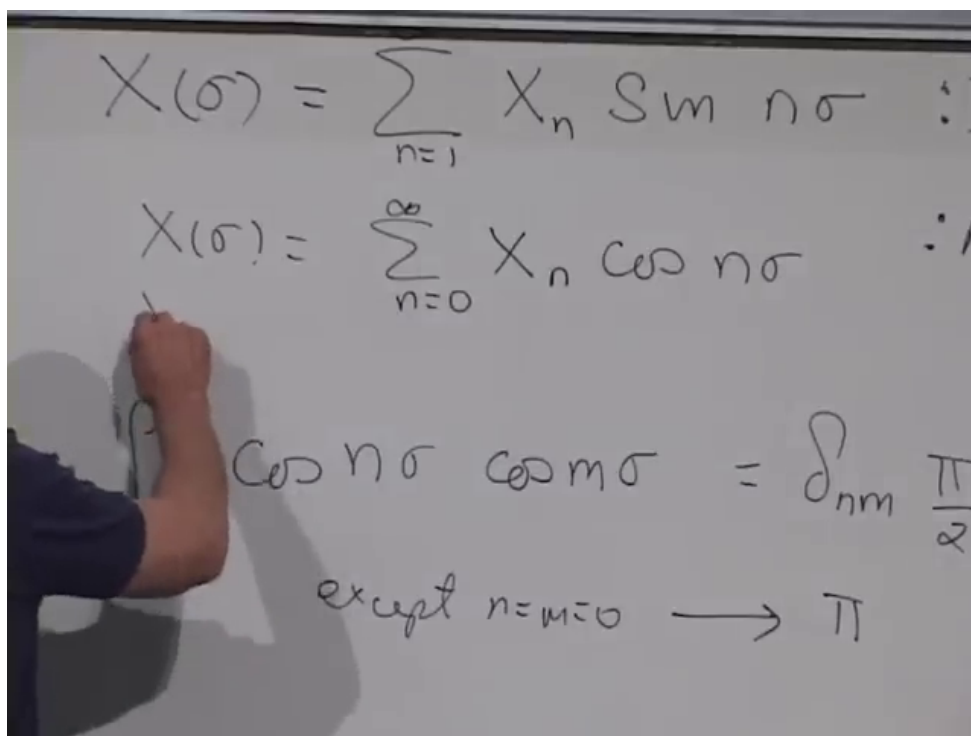


Figure 2.16: The Neumann mode expansion placed beside the orthogonality rule that will collapse the double sums.

Lecture frame: 01:21:45.

2.7.1 Kinetic term and the zero mode

Time differentiation gives

$$\partial_\tau X = \sum_{n=0}^{\infty} \dot{X}_n(\tau) \cos(n\sigma). \quad (2.35)$$

Squaring produces a double sum over n and m , but (2.10) removes every term with $n \neq m$. Treating the constant mode separately,

$$\frac{1}{2\pi} \int_0^\pi (\partial_\tau X)^2 d\sigma = \frac{1}{2} \dot{X}_0^2 + \frac{1}{4} \sum_{n=1}^\infty \dot{X}_n^2. \quad (2.36)$$

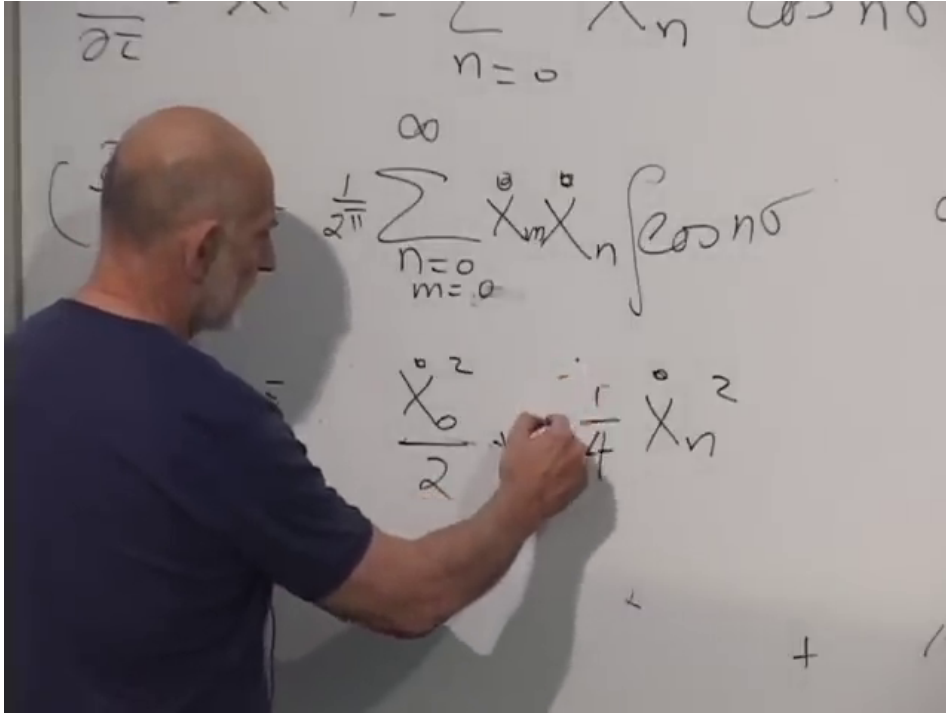


Figure 2.17: The orthogonality integral reducing the kinetic double sum to the zero mode plus a diagonal sum.

Lecture frame: 01:29:10.

The constant coefficient is the average position,

$$X_0(\tau) = \frac{1}{\pi} \int_0^\pi X(\sigma, \tau) d\sigma, \quad (2.37)$$

and similarly for Y_0 . With uniform analog mass density, these are the center-of-mass coordinates. They have kinetic energy but no restoring force: the entire string is free to move across the transverse plane.

2.7.2 Stretching term and oscillator frequencies

Spatial differentiation gives

$$\partial_\sigma X = - \sum_{n=1}^\infty n X_n(\tau) \sin(n\sigma). \quad (2.38)$$

The zero mode disappears because a uniform translation does not stretch the string. Using sine orthogonality,

$$\frac{1}{2\pi} \int_0^\pi (\partial_\sigma X)^2 d\sigma = \frac{1}{4} \sum_{n=1}^{\infty} n^2 X_n^2. \quad (2.39)$$

Handwritten notes on a chalkboard showing the derivation of the orthogonality condition for the string modes. The main equation is:
$$+ \frac{1}{2\pi} \sum_{nm} n m X_n X_m \int_0^\pi \sin n\sigma \sin m\sigma d\sigma$$
 Below it, there is a note: $+ (x \rightarrow y)$ and a calculation: $\int_0^\pi \sin n\sigma \sin m\sigma d\sigma = \frac{\pi}{2} \delta_{nm}$

Figure 2.18: Differentiating the cosine series produces the sine double sum; orthogonality again enforces $n = m$.

Lecture frame: 01:31:40.

The X -sector Lagrangian is therefore

$$L_X = \frac{1}{2} \dot{X}_0^2 + \frac{1}{4} \sum_{n=1}^{\infty} (\dot{X}_n^2 - n^2 X_n^2), \quad (2.40)$$

with an identical expression for Y . The corresponding energy has a plus sign in front of $n^2 X_n^2$. Every nonzero mode is an independent harmonic oscillator, and comparison with (2.14) gives

$$\omega_n = n, \quad n = 1, 2, 3, \dots \quad (2.41)$$

Handwritten notes on a chalkboard:

- At the top left, $n=0$ is written.
- Below it, $\omega_n = n$ is written.
- In the center, a large curly bracket groups the following expression:

$$\frac{\dot{X}_0^2}{2} + \left(\sum_n \frac{1}{4} \dot{X}_n^2 - \frac{1}{4} \sum_n n^2 X_n^2 \right)$$
- Below the bracketed expression, there is a plus sign and a right arrow pointing to y : $+ \quad \rightarrow y$
- At the bottom right, the expression $\omega = \pi/N$ is written, followed by $\frac{1}{N} =$.

Figure 2.19: The decoupled oscillator Lagrangian and the conclusion $\omega_n = n$, with the identical Y -tower indicated below.

Lecture frame: 01:35:50.

The normal modes are simply the harmonics of the string. The first bends slowly, the second has twice the frequency, the third three times, and so on. Orthogonality is what prevents one harmonic from coupling to another in the free theory.

Question from the audience. What happened to the $n = 0$ oscillator?^a

Response. It has zero frequency because it is not an internal oscillator at all. It translates the center of mass. No spring restores the whole string to a preferred transverse position, so the zero mode is a free particle coordinate.

^aLecture timestamp: 01:34:49.

Dropping the center-of-mass kinetic term leaves the internal energy

$$E_{\text{int}} = \frac{1}{4} \sum_{n=1}^{\infty} \left[\dot{X}_n^2 + n^2 X_n^2 + \dot{Y}_n^2 + n^2 Y_n^2 \right]. \quad (2.42)$$

In the light-front normalization this is proportional to the invariant mass squared of the string state. We have thus converted the relativistic mass problem into the quantum mechanics of two countable oscillator towers.

2.8 Quantized levels and particle behavior

Let $a_n^\dagger$ create one quantum in the X -oscillator of frequency n , and $b_n^\dagger$ do the same for Y . The lecture postpones a systematic operator treatment, but this notation makes the level counting transparent. Relative to the ground-state offset,

$$\Delta E_{\text{int}} = \sum_{n=1}^{\infty} n (Q_n^X + Q_n^Y), \quad (2.43)$$

$$Q_n^X, Q_n^Y \in \{0, 1, 2, \dots\}.$$

where $\hbar = 1$. The integer n labels which oscillator is excited; Q_n counts how many quanta occupy it.²

The oscillator vacuum $|0\rangle$ is not empty spacetime. It is one open string with no internal oscillator excited. It has energy, momentum, and a definite place in the string spectrum.

At level one there are two states,

$$a_1^\dagger |0\rangle, \quad b_1^\dagger |0\rangle, \quad (2.44)$$

corresponding to the lowest harmonic polarized along X or Y . Linear combinations describe any transverse polarization.

At level two there are five states:

$$(a_1^\dagger)^2 |0\rangle, \quad a_1^\dagger b_1^\dagger |0\rangle, \quad (b_1^\dagger)^2 |0\rangle, \quad a_2^\dagger |0\rangle, \quad b_2^\dagger |0\rangle. \quad (2.45)$$

Three use two quanta of the first harmonic; two use one quantum of the second harmonic. Both constructions cost two units.

²Editorial clarification: Each oscillator also contributes a zero-point energy. The infinite sum requires regularization and supplies the intercept of the string mass formula. The lecture deliberately leaves the ground-state offset unspecified here and counts only excitation levels.

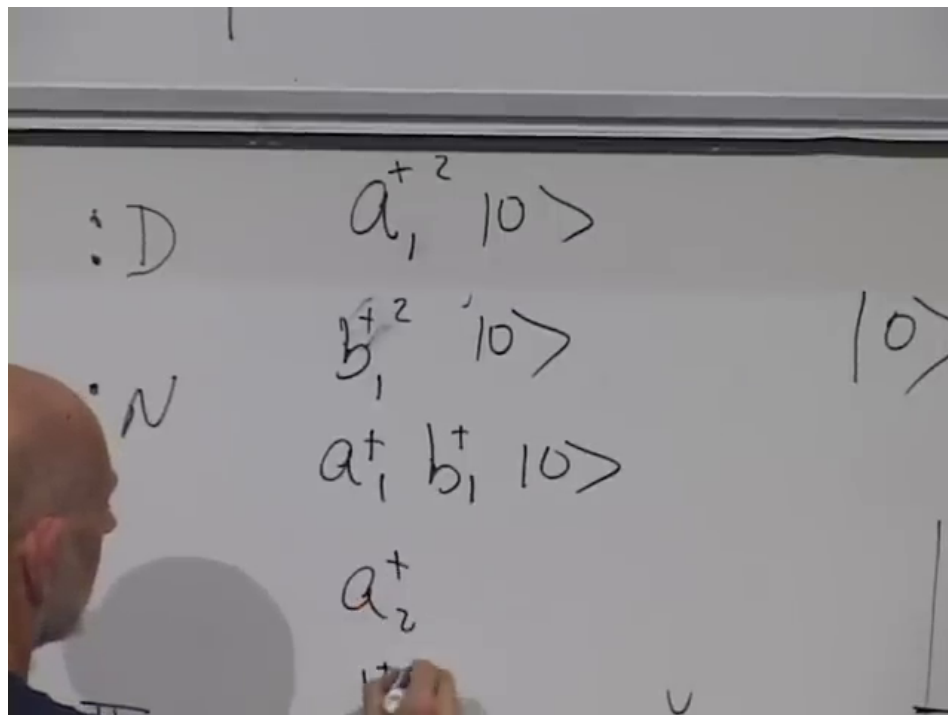


Figure 2.20: The five level-two states being recounted after an audience challenge: two double excitations, one mixed excitation, and two second-harmonic excitations.

Lecture frame: 01:44:00.

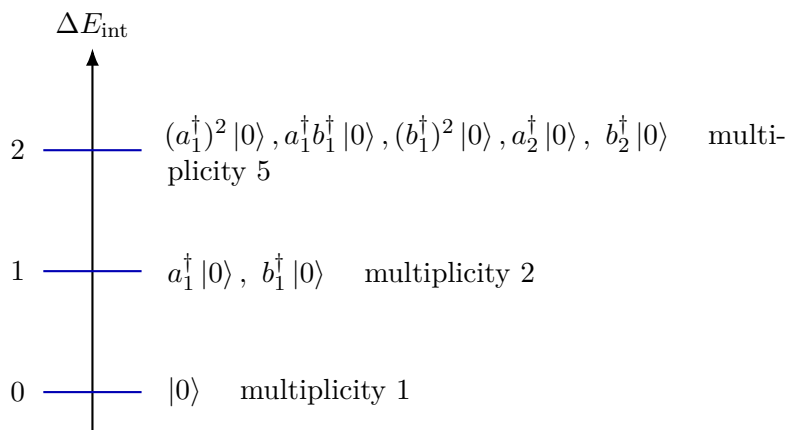


Figure 2.21: The first three excitation levels of the two-transverse-direction open string.

Question from the audience. Are there four or five states at the second excited level?^a

Response. Five. Exciting X_1 twice, exciting Y_1 twice, and exciting each once give three states. Exciting X_2 once or Y_2 once gives two more. The proposed exercise is to count the next two levels.

^aLecture timestamp: 01:43:04.

The essential conclusion is already visible. Successive levels are separated by finite gaps, and each has a finite multiplicity. The quantized string is therefore particle-like by the spectral criterion that

distinguished it from coffee, gas, or an ordinary violin string. Later we shall determine which states are physically welcome and which pathologies must be removed.

2.9 Why open-string interactions produce closed strings

An open string has two endpoints. Its elementary interactions join and split strings. In the hadronic picture one may imagine a quark at one end and an antiquark at the other; their annihilation joins two string segments. The fundamental strings of interest later need not carry that literal hadronic endpoint interpretation, but the local joining rule remains.

Now let one open string fluctuate until its own endpoints approach one another. Locally, an endpoint cannot know that the other endpoint belongs to the same string rather than a different one. If endpoints may join at all, they may join here, producing a closed loop.

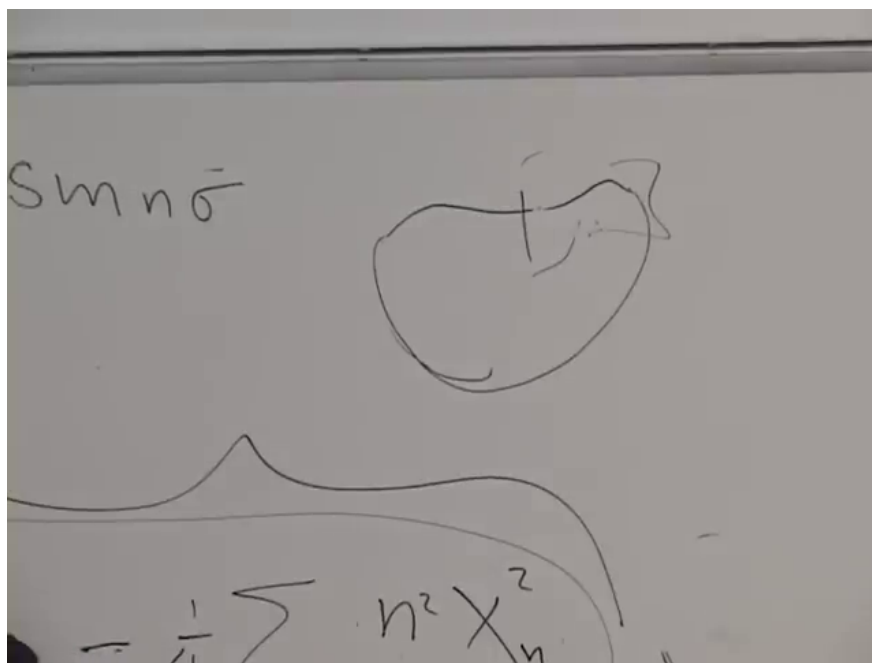


Figure 2.22: The board sketch of an open string fluctuating until its two endpoints meet and form a closed loop.

Lecture frame: 01:47:00.

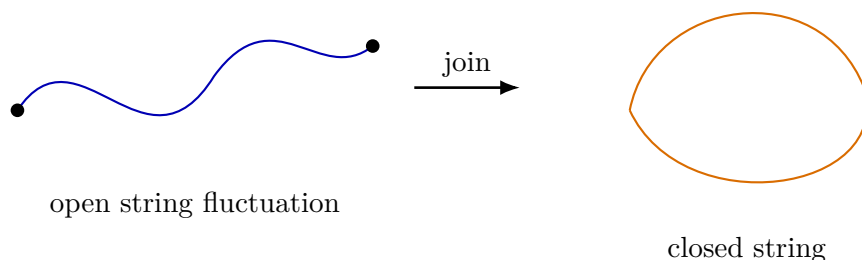


Figure 2.23: Once endpoint joining is allowed, an open string can close on itself.

Thus a theory may contain closed strings without open strings, but an interacting theory with open strings cannot consistently avoid closed strings. This asymmetry is the important prediction. The spectral identifications will be developed later, but the destination is already visible: massless open-string states often supply gauge particles such as photons, while the closed-string spectrum contains a massless spin-two state, the graviton. Closed strings, and therefore gravity, are not an optional decoration once string theory is interacting.

Chapter 3

From Open-String Quanta to Tachyons and Gravity

Overview. We now quantize the normal modes found in the previous lecture and ask what particle states they produce. Each transverse mode is a harmonic oscillator, but its normalization is slightly unfamiliar: exciting mode n raises the light-front energy, and hence the mass-squared level, by n units. The first excited open-string level transforms exactly like the two transverse polarizations of a massless vector. Relativistic consistency therefore fixes the ground-state mass squared to be negative. That state is the bosonic-string tachyon, whose real significance is an unstable vacuum, not a practical faster-than-light particle. Finally, once strings interact, joining and splitting make closed strings unavoidable; the lowest closed-string excitation will be the graviton.

3.1 The oscillator hidden in each string mode

For one transverse coordinate X , the n th nonzero normal mode has energy and Lagrangian

$$E_n = \frac{1}{4} \dot{X}_n^2 + \frac{n^2}{4} X_n^2, \quad (3.1)$$

$$L_n = \frac{1}{4} \dot{X}_n^2 - \frac{n^2}{4} X_n^2. \quad (3.2)$$

The factors of $1/4$ are not an error. They record the normalization chosen when the string action was reduced to independent Fourier modes. The first step in constructing the Hamiltonian is therefore to compute the canonical momentum,

$$p_n = \frac{\partial L_n}{\partial \dot{X}_n} = \frac{1}{2} \dot{X}_n, \quad \dot{X}_n = 2p_n. \quad (3.3)$$

The Legendre transform gives

$$H_n = p_n \dot{X}_n - L_n = p_n^2 + \frac{n^2}{4} X_n^2. \quad (3.4)$$

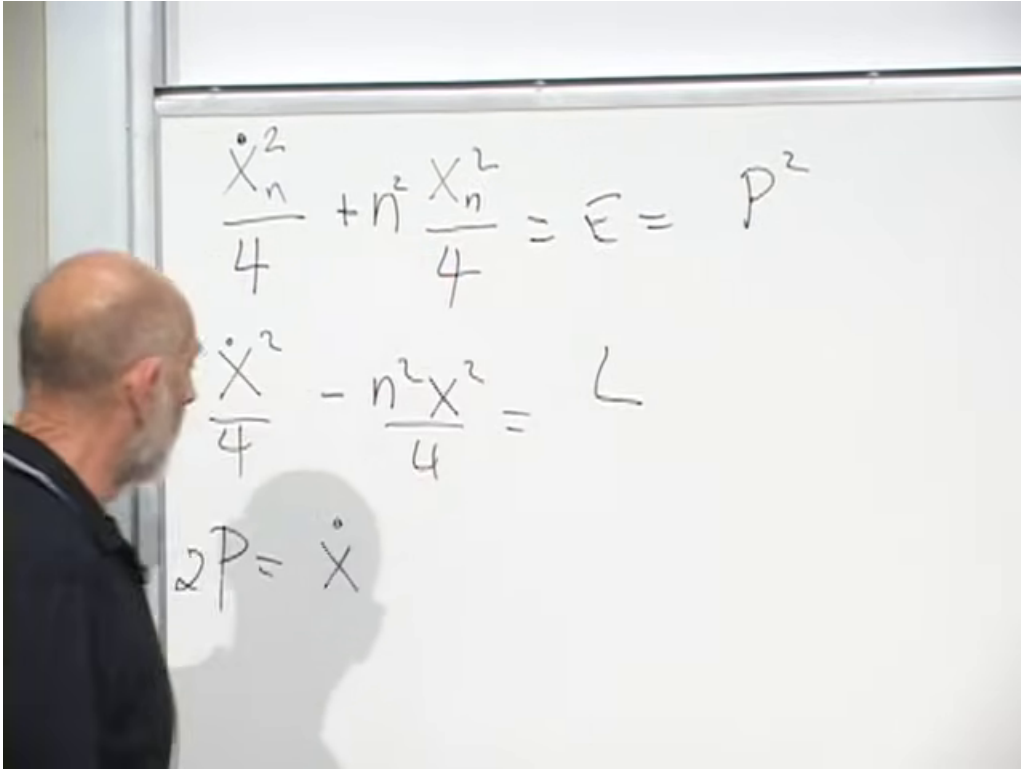


Figure 3.1: The mode energy, Lagrangian, canonical momentum, and Hamiltonian in the normalization inherited from the open string.

Lecture frame: 00:05:40.

To factorize H_n , combine X_n and p_n as

$$\frac{nX_n}{2} \pm ip_n. \quad (3.5)$$

Classically the two factors would commute, but quantum mechanically

$$[X_n, p_m] = i\delta_{nm} \quad (3.6)$$

in units with $\hbar = 1$. The cross terms do not cancel innocently. For a fixed mode,

$$\left[\frac{nX_n}{2} + ip_n, \frac{nX_n}{2} - ip_n \right] = n, \quad (3.7)$$

$$\left(\frac{nX_n}{2} - ip_n \right) \left(\frac{nX_n}{2} + ip_n \right) = H_n - \frac{n}{2}. \quad (3.8)$$

The image shows a whiteboard with handwritten mathematical expressions. The top expression is $= \left[p^2 + \frac{n^2 x^2}{4} \right] = H$. Below it, a commutator is calculated: $\left[\left(\frac{nx}{2} + ip \right) \left(\frac{nx}{2} - ip \right) \right] = \frac{n}{2} (-i)(i) = n$. Arrows indicate the expansion of the commutator into two terms.

Figure 3.2: The board calculation in which the canonical commutator supplies the factor of n .

Lecture frame: 00:10:30.

The factor of n tells us how to normalize the ladder operators:

$$a_n \equiv \frac{\sqrt{n}}{2} X_n + \frac{i}{\sqrt{n}} p_n, \quad (3.9)$$

$$a_n^\dagger \equiv \frac{\sqrt{n}}{2} X_n - \frac{i}{\sqrt{n}} p_n. \quad (3.10)$$

They obey

$$[a_n, a_m^\dagger] = \delta_{nm}, \quad H_n = n \left(a_n^\dagger a_n + \frac{1}{2} \right). \quad (3.11)$$

Thus $a_n^\dagger$ adds one quantum of the n th oscillator, but that quantum carries n units of mode energy. The familiar oscillator spacing is not lost; its frequency is $\omega_n = n$.

Handwritten notes on a whiteboard:

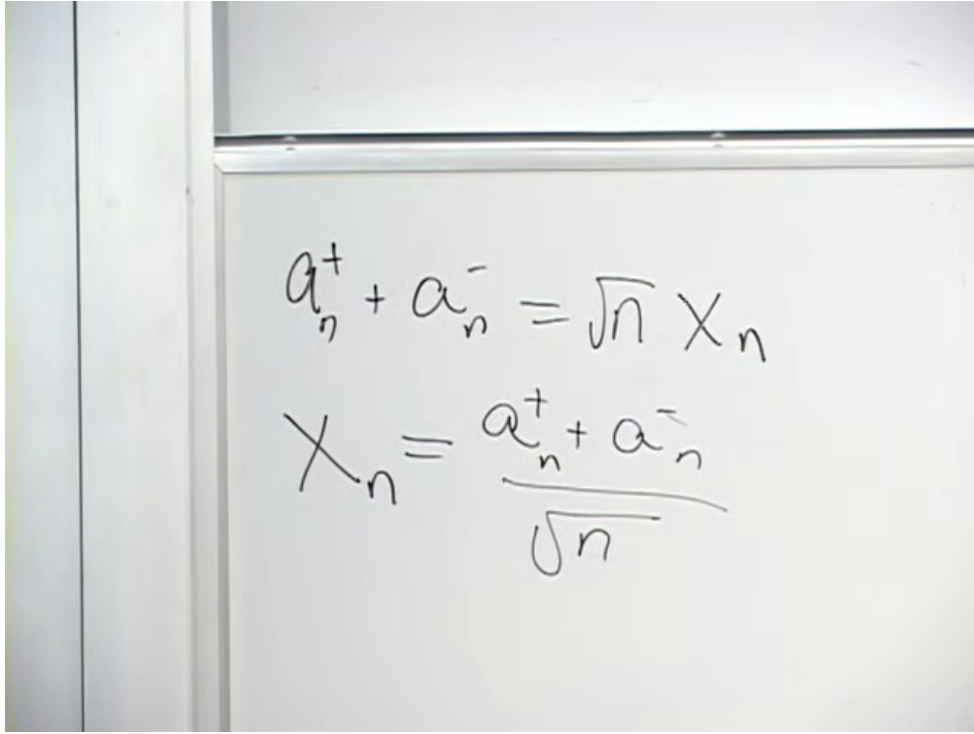
- At the top: $[X, P] = i$
- To the right, partially visible: $\frac{1}{\sqrt{n}} \left(\frac{nX}{2} - \dots \right)$
- Inside a hand-drawn box:
 - $a^- = \frac{\sqrt{n} X}{2} + \frac{i}{\sqrt{n}} P$
 - $a^+ = \frac{\sqrt{n} X}{2} - \frac{i}{\sqrt{n}} P$

Figure 3.3: Normalized annihilation and creation operators for the n th string oscillator.*Lecture frame: 00:12:00.*

3.1.1 Two transverse towers

In four spacetime dimensions, a string carrying very large momentum along the z -axis has two transverse coordinates, X and Y . We therefore need two independent oscillator towers. Let $a_n, a_n^\dagger$ act on the X modes and $b_n, b_n^\dagger$ on the Y modes. Solving (3.9)–(3.10) for the coordinates gives

$$X_n = \frac{a_n + a_n^\dagger}{\sqrt{n}}, \quad Y_n = \frac{b_n + b_n^\dagger}{\sqrt{n}}. \quad (3.12)$$



$$a_n^+ + a_n^- = \sqrt{n} X_n$$

$$X_n = \frac{a_n^+ + a_n^-}{\sqrt{n}}$$

Figure 3.4: Solving for a mode coordinate in terms of its creation and annihilation operators.

Lecture frame: 00:17:05.

The open-string coordinates were expanded in Neumann modes,

$$X(\sigma, \tau) = X_{\text{cm}}(\tau) + \sum_{n=1}^{\infty} X_n(\tau) \cos(n\sigma), \quad (3.13)$$

$$Y(\sigma, \tau) = Y_{\text{cm}}(\tau) + \sum_{n=1}^{\infty} Y_n(\tau) \cos(n\sigma). \quad (3.14)$$

Cosines appear because the endpoint slopes vanish. Substituting (3.12) turns the string itself into an operator:

$$X(\sigma, \tau) = X_{\text{cm}}(\tau) + \sum_{n=1}^{\infty} \frac{a_n(\tau) + a_n^\dagger(\tau)}{\sqrt{n}} \cos(n\sigma), \quad (3.15)$$

$$Y(\sigma, \tau) = Y_{\text{cm}}(\tau) + \sum_{n=1}^{\infty} \frac{b_n(\tau) + b_n^\dagger(\tau)}{\sqrt{n}} \cos(n\sigma). \quad (3.16)$$

$$X_n = \frac{a_n + \bar{a}_n}{\sqrt{n}}, \quad Y_n = \frac{b_n + \bar{b}_n}{\sqrt{n}}$$

$$X(\sigma) = \sum_n X_n \cos n\sigma$$

$$Y(\sigma) = \sum_n Y_n \cos n\sigma$$

(a) Coordinate-mode expansions.

$$X_n = \frac{a_n + \bar{a}_n}{\sqrt{n}}$$

$$Y_n = \frac{b_n + \bar{b}_n}{\sqrt{n}}$$

$$X(\sigma) = \sum_n \frac{a_n + \bar{a}_n}{\sqrt{n}} \cos n\sigma$$

$$Y(\sigma) = \sum_n \frac{b_n + \bar{b}_n}{\sqrt{n}} \cos n\sigma$$

(b) Ladder operators substituted into the string coordinates.

Figure 3.5: The shape of the quantum open string is assembled from all of its transverse oscillator operators.

Lecture frame: 00:20:35 and 00:21:05.

Question from the audience. Is X_n also a function of σ ?^a

Response. No. The σ -dependence is entirely in $\cos(n\sigma)$. The coefficient X_n is the time-dependent amplitude of that spatial pattern, so $X_n = X_n(\tau)$.

^aLecture timestamp: 00:21:48.

Question from the audience. Why do the operators called raising and lowering seem to have their signs reversed?^a

Response. The visible sign of ip_n is not a naming rule. Compute the commutator with the Hamiltonian. The operator satisfying $[H_n, a_n^\dagger] = na_n^\dagger$ raises the energy; the one satisfying $[H_n, a_n] = -na_n$ lowers it.

^aLecture timestamp: 00:23:13.

The two transverse towers are a pedagogical four-dimensional picture. A bosonic string quantized consistently has twenty-six spacetime dimensions, so it has twenty-four transverse oscillator families rather than two. The logic below is unchanged: the number of first-level states must match the number of physical polarizations of a massless vector in the same dimension.

3.2 Massive spin and massless helicity

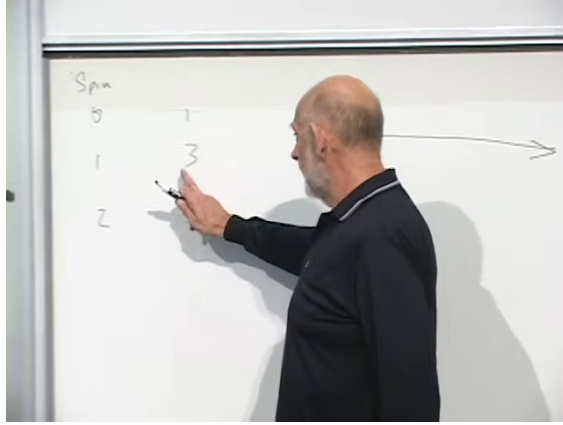
Before assigning spin to the string states, we need one sharp distinction. A massive particle of spin J has $2J + 1$ spin states,

$$m_J = -J, -J + 1, \dots, J - 1, J. \quad (3.17)$$

For example, a massive spin-one particle has three polarizations and a massive spin-two particle has five. We can see why by going to the rest frame. There every spatial rotation is available. Rotate

the spin axis, boost back, and all $2J + 1$ components belong to the same relativistic particle.

A massless particle has no rest frame. Its invariant spin label is helicity, the angular momentum along its direction of motion. A photon has only $h = +1$ and $h = -1$; a graviton has only $h = +2$ and $h = -2$. A scalar has one $h = 0$ state. The lecture uses reflection symmetry to pair the two helicities. In a chiral theory a single massless irreducible representation may carry only one helicity, so the pairing is an additional symmetry statement rather than a consequence of masslessness alone.¹



(a) The spin-state count begun on the board.

Spin	Massive states	Massless helicities
0	1	0
1	3	+1, -1
2	5	+2, -2

Figure 3.6: Mass removes the rest frame and with it the intermediate spin projections.

Lecture frame: 00:27:40.

3.2.1 Linear and circular photon polarization

Take a photon moving along z . Its two transverse linear polarizations are $|x\rangle$ and $|y\rangle$. They are not helicity eigenstates. The circular states are

$$|R\rangle = \frac{1}{\sqrt{2}} (|x\rangle + i|y\rangle), \quad (3.18)$$

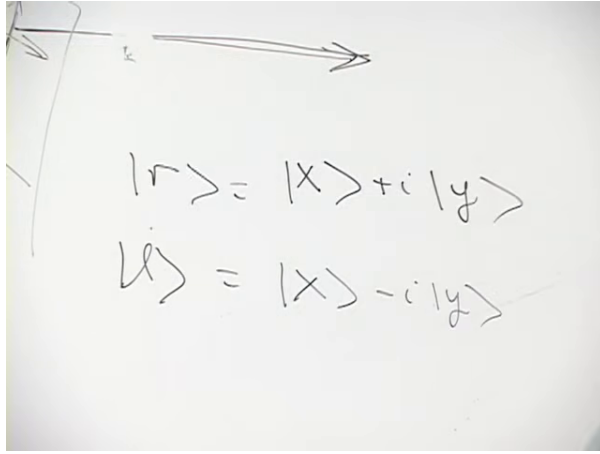
$$|L\rangle = \frac{1}{\sqrt{2}} (|x\rangle - i|y\rangle), \quad (3.19)$$

with helicities $+1$ and -1 , up to the convention used to name right and left. A linearly polarized photon is a superposition of those two outcomes. Its average helicity can vanish, but a helicity measurement never finds a spin-zero photon. Likewise,

$$\frac{1}{\sqrt{2}} (|x\rangle + |y\rangle) \quad (3.20)$$

is simply linear polarization at forty-five degrees.

¹Editorial clarification: The modern little-group classification permits one-helicity massless representations. Parity or an appropriate CPT-complete field content supplies the opposite helicity in the familiar photon and graviton examples.



(a) Circular polarization states on the board.

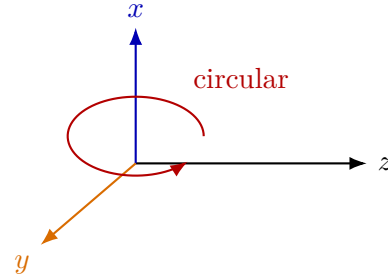


Figure 3.7: Linear polarization supplies a transverse basis; circular combinations diagonalize helicity.

Lecture frame: 00:37:45.

Question from the audience. Is this like the separation of spin states in a Stern–Gerlach experiment?^a

Response. The quantum-superposition idea is the same, but for the photon the direct language is polarization. A circular-polarization measurement distinguishes the two helicities; there is no longitudinal spin-zero photon in the beam.

^aLecture timestamp: 00:37:40.

The graviton repeats the pattern at spin two. For spin bookkeeping one may write schematically

$$|+2\rangle \sim |+1\rangle \otimes |+1\rangle, \quad |-2\rangle \sim |-1\rangle \otimes |-1\rangle. \quad (3.21)$$

This does not say that a graviton is literally made from two photons. It says that two equal helicities add to the correct spin-two helicity.

Question from the audience. Does the two-photon analogy extend to other quantum numbers?^a

Response. Not as a constituent model. There are powerful formulas in which gravity amplitudes look like products, or “squares,” of gauge-theory amplitudes, but the graviton is still its own particle. The analogy is most trustworthy here as a statement about helicity.^b

^aLecture timestamp: 00:41:05.^bEditorial clarification: Modern Kawai–Lewellen–Tye and double-copy relations make the amplitude statement precise in suitable theories; they do not turn a graviton into a bound state of two photons.

Question from the audience. A photon can have any momentum, so does it not have a continuous spectrum rather than the isolated spectrum required of a particle?^a

Response. The continuous variable is the center-of-mass momentum. Hold that momentum fixed and inspect the internal invariant mass. A particle is identified by an isolated mass level; varying its overall momentum then produces the usual continuous dispersion relation.

^aLecture timestamp: 00:42:17.

3.3 The lowest open-string levels

We can now return to the string. In the large-momentum description, the finite internal light-front energy is proportional to invariant mass squared. Suppressing the overall string scale, write the spectrum as

$$M^2 = m_0^2 + \sum_{n=1}^{\infty} n \left(N_n^{(a)} + N_n^{(b)} \right), \quad (3.22)$$

$$N_n^{(a)} = a_n^\dagger a_n, \quad N_n^{(b)} = b_n^\dagger b_n.$$

The oscillator ground state obeys

$$a_n |0\rangle = 0, \quad b_n |0\rangle = 0, \quad n = 1, 2, \dots \quad (3.23)$$

It is not an empty absence of string. It is the least excited quantum string, and its coordinates still fluctuate. Let its mass squared be m_0^2 , not yet determined.

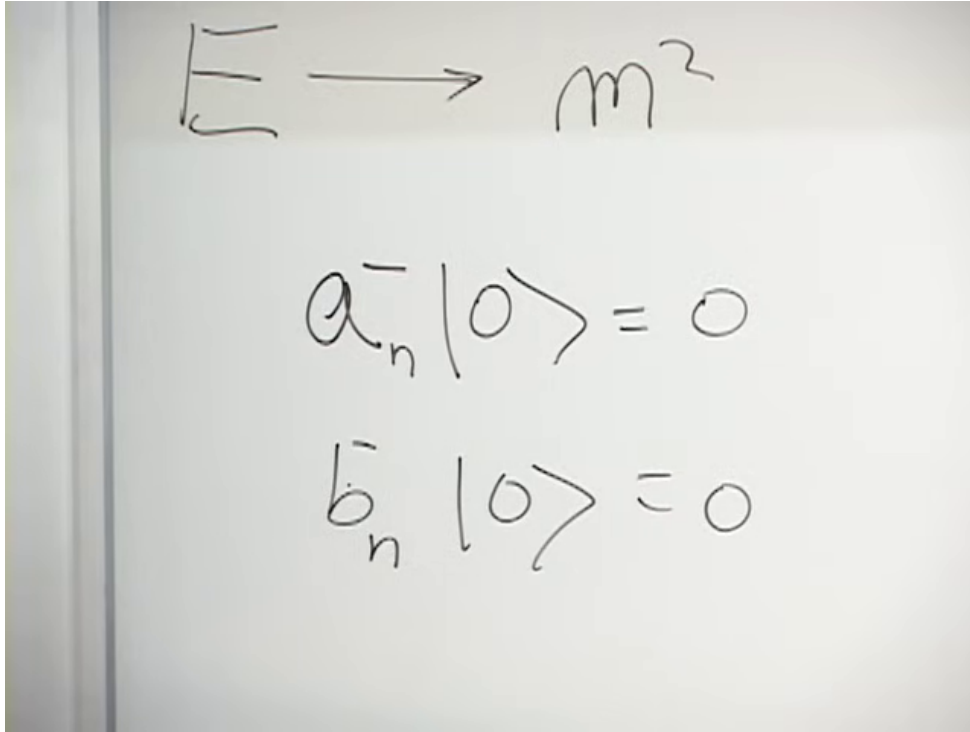


Figure 3.8: The open-string ground state and the annihilation conditions that define it.

Lecture frame: 00:45:20.

The cheapest excitations use the $n = 1$ creation operators:

$$a_1^\dagger |0\rangle, \quad b_1^\dagger |0\rangle, \quad M^2 = m_0^2 + 1. \quad (3.24)$$

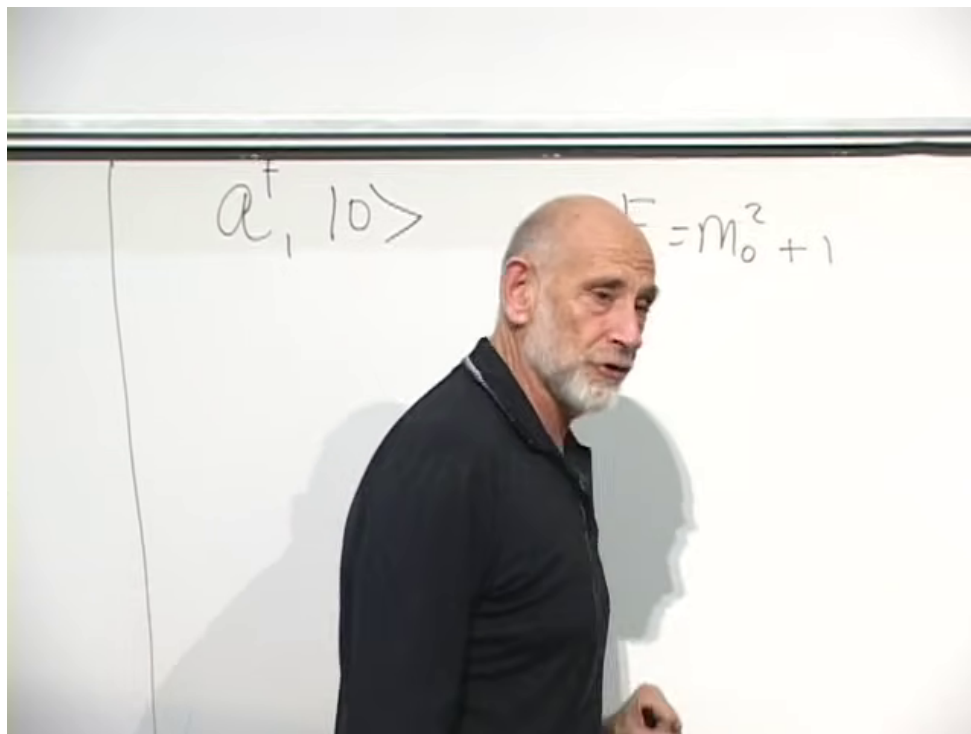


Figure 3.9: One of the two degenerate first-level states, with its mass-squared increment written on the board.

Lecture frame: 00:49:25.

The “one” in [eqrefeq:l03-first-level](#) is not one joule, one electron volt, or one Planck mass. We have suppressed the dimensionful string scale. For now it means one step in the oscillator’s mass-squared ladder; the missing units will be restored in the next lecture.

Any linear combination of the two degenerate states is also a state at this level. In particular,

$$\frac{1}{\sqrt{2}} \left(a_1^\dagger \pm i b_1^\dagger \right) |0\rangle \quad (3.25)$$

is the string analogue of circular polarization. Under a rotation in the x - y plane, $a_1^\dagger$ and $b_1^\dagger$ transform as the two components of a transverse vector. The combinations in [\(3.25\)](#) therefore have helicity $+1$ and -1 .

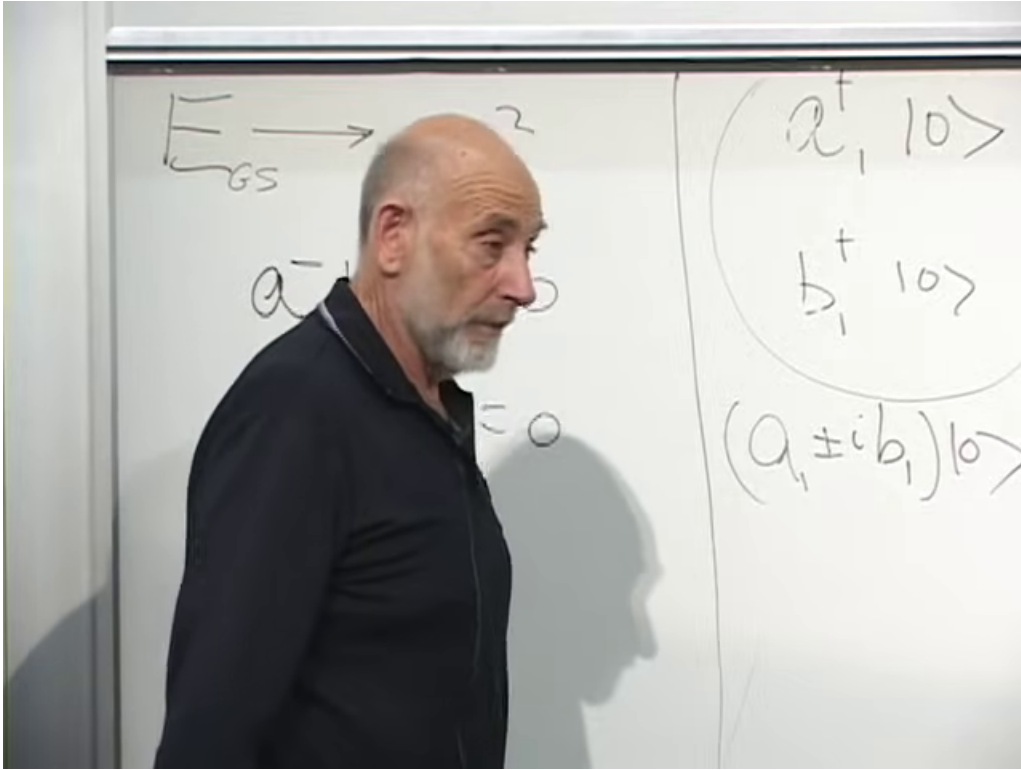


Figure 3.10: Ground-state conditions, the two first-level excitations, and their circular combinations collected on the board.

Lecture frame: 00:56:00.

There is no third degenerate state that could serve as a longitudinal polarization. If these two states had nonzero mass, a spin-one relativistic particle would require three polarizations. With exactly two, the state must be massless. Hence

$$m_0^2 + 1 = 0, \quad m_0^2 = -1. \quad (3.26)$$

The first level is photon-like: it is a massless gauge-vector state. It need not be the photon of electromagnetism; its eventual gauge identity depends on the string model. The ground state below it has negative mass squared. That is the bosonic open-string tachyon.²

Question from the audience. Why should we care that m_0^2 is negative?^a

Response. Because mass squared is the curvature of the quadratic field potential about the state we have called the vacuum. Negative curvature means that this configuration is not a stable ground state. The next calculation makes that statement precise.

^aLecture timestamp: 00:56:35.

²Editorial clarification: In the full critical bosonic string the normal-ordering intercept and Lorentz algebra are tied to $D = 26$. The lecture's two-transverse-direction calculation is the four-dimensional shadow of the same state-counting argument: $D - 2$ first-level oscillator states match the $D - 2$ physical polarizations of a massless vector.

3.4 What a tachyon really means

The name “tachyon” invites a misleading first thought. Start from the relativistic dispersion relation with $c = 1$,

$$E(p) = \sqrt{p^2 + m^2}. \quad (3.27)$$

For a wave packet, the group velocity is

$$v_g = \frac{dE}{dp}. \quad (3.28)$$

This reproduces familiar cases:

$$E_{\text{nr}}(p) = \frac{p^2}{2m}, \quad v_g = \frac{p}{m}, \quad (3.29)$$

$$E_\gamma(p) = |p|, \quad |v_g| = 1, \quad (3.30)$$

$$E_m(p) = \sqrt{p^2 + m^2}, \quad |v_g| = \frac{|p|}{\sqrt{p^2 + m^2}} < 1. \quad (3.31)$$

Handwritten equations on a whiteboard:

$$E(p) = \frac{p^2}{2m}$$

$$\frac{\partial E}{\partial p} = \frac{p}{m} = v$$

(a) Nonrelativistic group velocity.

Handwritten equations on a whiteboard:

$$E = \sqrt{p^2 + m^2}$$

$$\frac{\partial E}{\partial p} = \frac{p}{\sqrt{p^2 + m^2}}$$

(b) Relativistic dispersion and its derivative.

Figure 3.11: Reading propagation speed from the slope of $E(p)$.

Lecture frame: 01:01:50 and 01:03:15.

If $m^2 = -\mu^2$, the formal expression becomes

$$E(p) = \sqrt{p^2 - \mu^2}, \quad v_g = \frac{p}{\sqrt{p^2 - \mu^2}}, \quad (3.32)$$

which has $|v_g| > 1$ whenever the square root is real. But this is not the right physical interpretation of the bosonic-string ground state. The trouble appears already at low momentum, where the frequency is imaginary.

3.4.1 Negative curvature of the field potential

Consider a real scalar field in one spatial dimension,

$$\left(\partial_t^2 - \partial_x^2 + m^2\right)\phi = 0. \quad (3.33)$$

For a plane wave $\phi \propto e^{-i\omega t + ikx}$,

$$\omega^2 = k^2 + m^2. \quad (3.34)$$

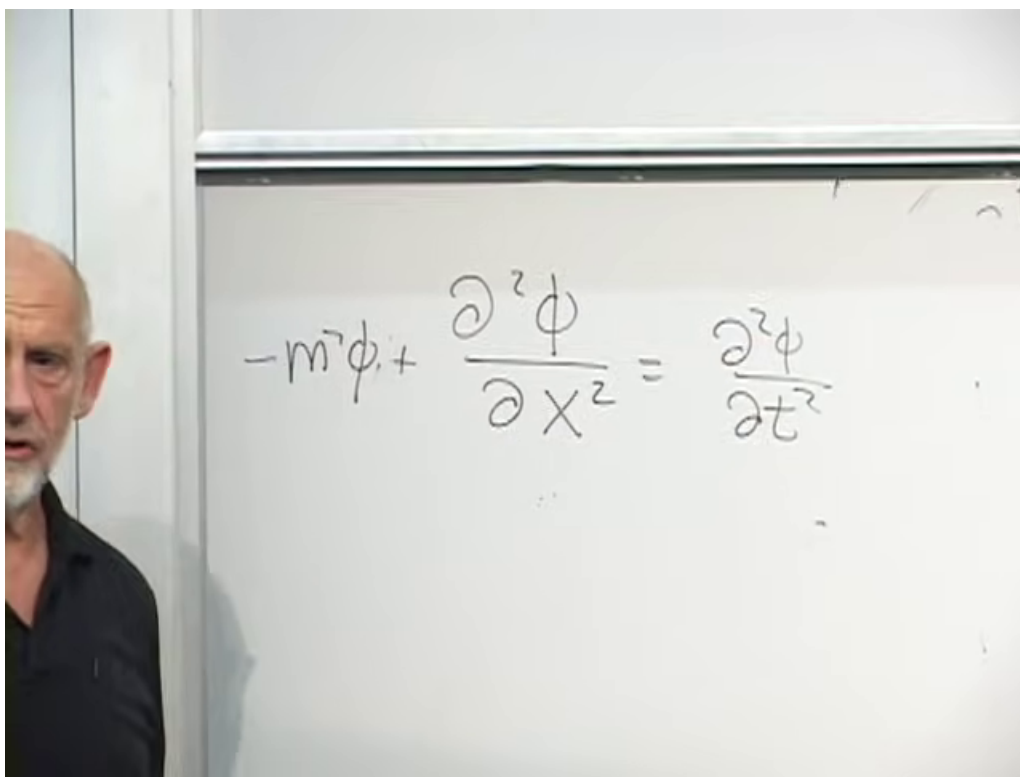


Figure 3.12: The Klein–Gordon equation used to expose the instability behind negative mass squared.

Lecture frame: 01:06:40.

The conserved quadratic energy is

$$\mathcal{E} = \frac{1}{2} \int dx \left[\dot{\phi}^2 + (\partial_x \phi)^2 + m^2 \phi^2 \right]. \quad (3.35)$$

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{\partial^2 \phi}{\partial t^2}$$

$$\frac{1}{2} \left(\frac{\partial \phi}{\partial t} \right)^2 + \frac{1}{2} \left(\frac{\partial \phi}{\partial x} \right)^2 + \frac{m^2 \phi^2}{2}$$

Figure 3.13: The scalar-field energy, with the mass term supplying the local curvature of the potential.

Lecture frame: 01:08:10.

For $m^2 > 0$, the local potential $V(\phi) = \frac{1}{2}m^2\phi^2$ is a bowl. The configuration $\phi = 0$ sits at its bottom. For $m^2 = -\mu^2$, the same point is the top of an inverted bowl. Equation (3.34) becomes

$$\omega^2 = k^2 - \mu^2. \quad (3.36)$$

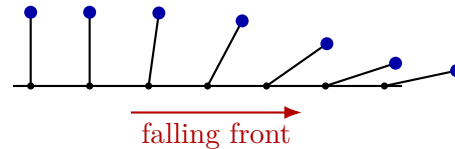
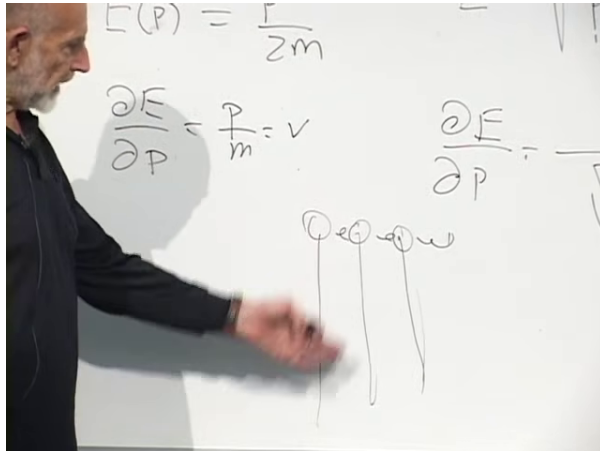
Modes with $k^2 < \mu^2$ have

$$\omega = \pm i\sqrt{\mu^2 - k^2}, \quad \phi_k(t) \propto e^{\pm\sqrt{\mu^2 - k^2}t}. \quad (3.37)$$

One branch grows exponentially. The alleged vacuum is therefore unstable: an arbitrarily small long-wavelength disturbance drives the field away from $\phi = 0$. That is what a tachyon in the spectrum means.

3.4.2 The upright-pendulum chain

A row of coupled pendulums makes the distinction visible. Put every pendulum upright. The uniform configuration exists, but it is balanced at a maximum of potential energy. Nudge one pendulum and it falls, pulling on its neighbors. A front of falling pendulums spreads through the chain at the finite signal speed fixed by the couplings.



(a) The coupled upright-pendulum model on the board.

Figure 3.14: Instability spreads causally through a system even though its linearized spectrum contains negative mass squared.

Lecture frame: 01:10:40.

Nothing in this mechanism permits a controlled signal to outrun the causal front. A tachyonic mode diagnoses a bad expansion point, not a stable superluminal messenger.³

Question from the audience. Is the massless state just found the graviton predicted by string theory?^a

Response. No. It is the two-polarization spin-one state of an open string, so it is gauge-boson-like. The graviton is a spin-two excitation of a closed string.

^aLecture timestamp: 01:14:27.

Question from the audience. Is an upside-down potential mathematically allowed, and where would the system finish after it falls?^a

Response. It is a perfectly legitimate local approximation, but by itself it has no ground state because the energy is unbounded below. A more complete potential may bend upward again and possess a lower minimum. The system then rolls or condenses toward that stable configuration.

^aLecture timestamp: 01:15:17.

The ordinary bosonic string has precisely this defect. Its higher oscillator states remain part of the spectrum, but they are excitations of an unstable background. Consistent superstring theories remove the open-string tachyon in their stable sectors while retaining massless and massive string excitations.

³Editorial clarification: For a local relativistic field equation the front velocity remains the causal speed. The formal group velocity of the real high- k branch does not turn the unstable vacuum into a faster-than-light communication channel.

Question from the audience. Do the other properties of the open-string state match a photon as well as its two polarizations do?^a

Response. The state behaves as a massless gauge boson. Its detailed charge assignments, radiation rules, and identity depend on the string model, so one should call it photon-like rather than identify it immediately with electromagnetism. Those interaction properties are developed later.

^aLecture timestamp: 01:17:07.

3.5 How strings interact

The free spectrum is only half the theory. Strings also join and split. It is useful first to remember that an excited free string can carry angular momentum. A straight rotating string turns about its center of mass, and its energy and angular momentum rise together along the familiar stringy trajectory.

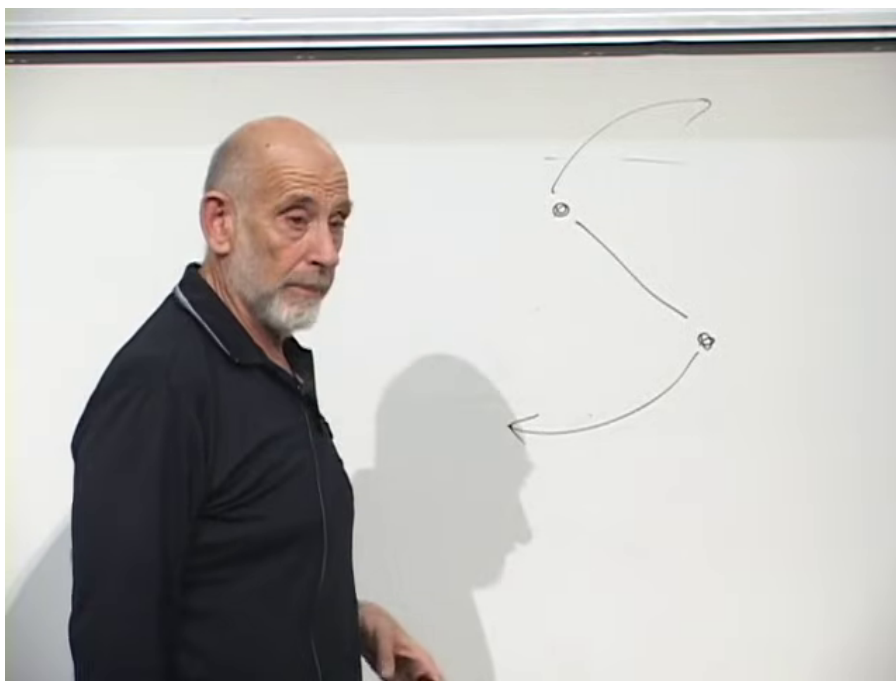


Figure 3.15: A rotating open string sketched about its center-of-mass point.

Lecture frame: 01:19:30.

Question from the audience. Can a free string rotate end over end about one fixed endpoint?^a

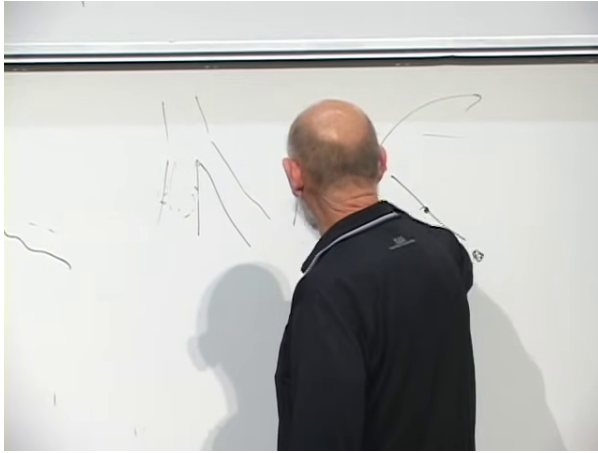
Response. Not without an anchor. A free isolated system rotates about its center of mass. If an endpoint is attached to another object or to a brane, the combined system can rotate about a different center, just as an electron and nucleus orbit their common center of mass while the much heavier nucleus appears nearly fixed.

^aLecture timestamp: 01:20:02.

3.5.1 The open-string vertex

For open strings the elementary interaction occurs when two endpoints meet at the same spacetime point. There is an amplitude g_s for them to join and form one string. Reversing the history gives an amplitude for one string to split into two. Schematically,

$$\text{open} + \text{open} \quad \rightleftharpoons \quad \text{open}. \quad (3.38)$$



(a) The joining history on the board.

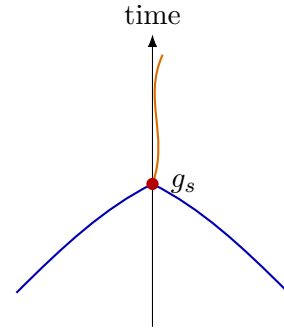


Figure 3.16: Endpoint joining is the string analogue of a local particle interaction vertex; time reversal gives splitting.

Lecture frame: 01:24:05.

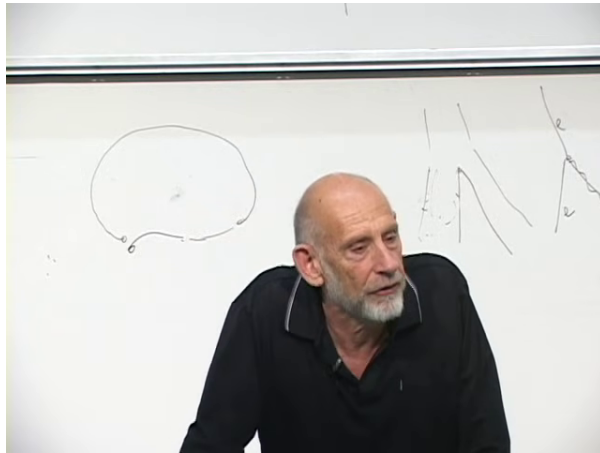
The coupling is an amplitude. For one specified elementary process the corresponding probability begins at order $|g_s|^2$, with phase-space and normalization factors supplied by the full calculation. Calling g_s a “probability” is therefore informal shorthand; keeping amplitude and probability distinct prevents a spurious square.

The interaction does not depend on whether the external string happens to be vibrating, rotating, or in some other oscillator state. Those states change the detailed matrix element, but the same local joining rule is used.

3.5.2 An open string can close itself

Suppose the two endpoints of one open string meet. The joining rule cannot know that they belong to the same string. They can fuse, producing a closed loop:

$$\text{open} \longrightarrow \text{closed}. \quad (3.39)$$



(a) The open string brings its endpoints together.

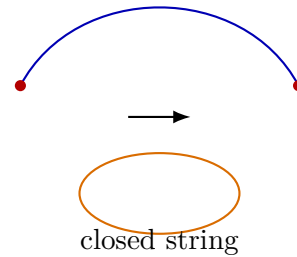


Figure 3.17: Interacting open strings inevitably generate closed strings.

Lecture frame: 01:27:20.

This is not an optional embellishment. Any theory of interacting open strings must contain closed-string states. The reverse process lets a closed string open into an open string when the allowed endpoints and boundary conditions are present.

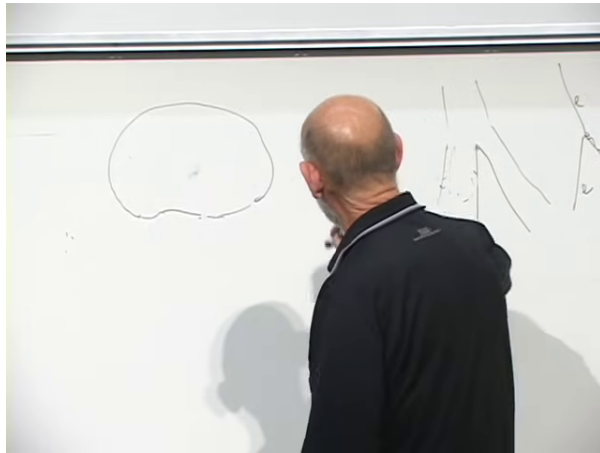
Question from the audience. When strings fuse or split, does their oscillator energy change, and does the theory determine how?^a

Response. Yes. Energy and momentum are conserved for the complete process, while excitation can be redistributed among the outgoing strings. The interaction vertex and the external wavefunctions determine the transition amplitude.

^aLecture timestamp: 01:28:05.

3.5.3 Closed strings can interact without opening

The implication is one-way. Interacting open strings require closed strings, but a theory may begin with only closed strings. Two closed loops can touch, reconnect, and become one closed loop; the reverse history splits one loop into two. No open endpoint need appear.



(a) Closed-string fusion on the board.

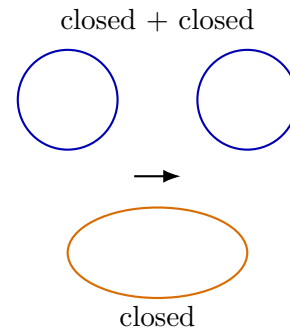


Figure 3.18: Closed strings possess their own joining and splitting interaction and need not pass through an open-string state.

Lecture frame: 01:29:05.

Question from the audience. If a theory starts with closed strings, must it also contain open strings?^a

Response. No. Closed strings can interact by reconnection while remaining closed throughout. Open strings require additional boundary data, such as allowed endpoints or branes.

^aLecture timestamp: 01:28:35.

3.6 Why gravity cannot be avoided

The closed-string oscillator calculation parallels the open-string one, but there are independent left-moving and right-moving excitations. Their lowest nontrivial matched level contains a massless spin-two state. That state has the helicities of a graviton. Consequently every consistent interacting closed-string theory contains gravity.

Question from the audience. Why does the lowest closed-string excitation have to be a graviton, just as the open-string level had to be photon-like? ^a

Response. The answer is the same kind of state counting, now with left- and right-moving oscillator vectors. Their tensor product contains a massless symmetric traceless spin-two polarization. The explicit closed-string spectrum will be worked out next; at this stage the point is the inevitability of that state.

^aLecture timestamp: 01:30:59.

Open and closed strings may also interact with one another. A closed loop can meet an open string and be absorbed, leaving a larger or differently excited open string. Conversely an open string can radiate a closed loop. At the particle level this is the emission or absorption of a graviton by matter.

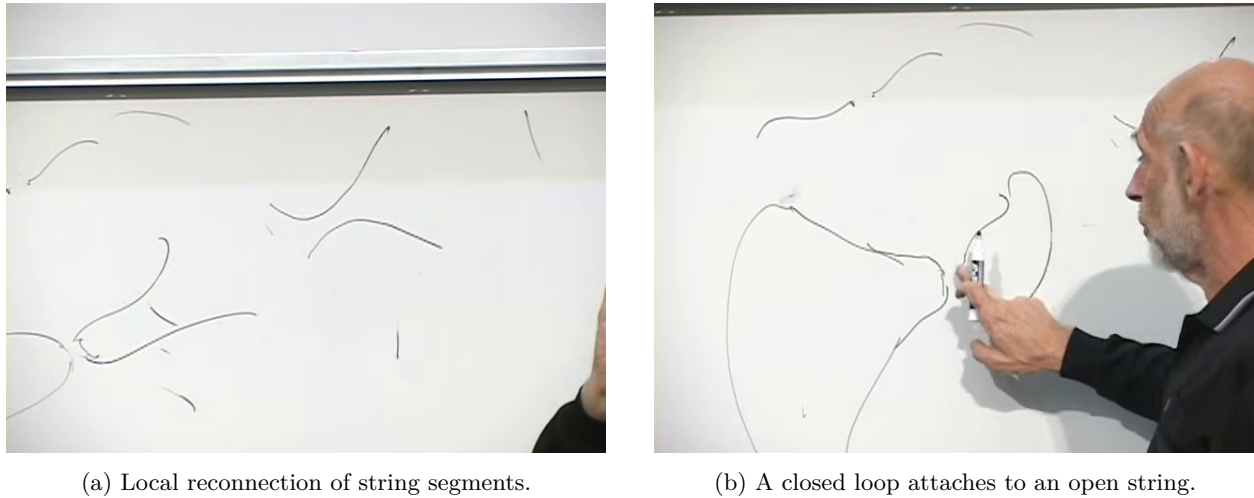


Figure 3.19: The board pictures behind mixed open–closed interactions and the universal possibility of graviton exchange.

Lecture frame: 01:33:10 and 01:34:35.

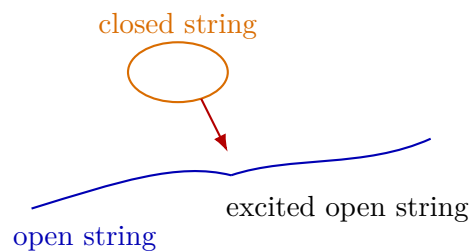


Figure 3.20: A clean spacetime-free schematic of closed-string absorption by an open string.

The graviton can couple to every string state because every string carries energy and momentum and the joining rules allow closed-string exchange. This is the string-theory origin of gravity’s universality.

Question from the audience. Can a closed string become an open string, or can it join an open string without violating the endpoint rule?^a

Response. In a theory containing both sectors, a closed string can attach to an open string and change its state. Whether a lone closed string may simply open depends on whether the theory supplies permissible endpoints. The local reconnection rule remains the same.

^aLecture timestamp: 01:31:20.

Question from the audience. Can closed loops become permanently linked or knotted?^a

Response. At weak coupling strings are mostly transparent: they can pass through one another unless the small reconnection amplitude acts. A link or knot can therefore persist for some time, but topology can change when an allowed string interaction occurs.

^aLecture timestamp: 01:35:10.

Question from the audience. Is the interaction probability independent of the state or speed of the string?^a

Response. The local coupling is the same parameter, but the total transition probability also depends on the wavefunctions and on how long the relevant pieces remain in contact. A fast encounter generally offers less interaction time than a slow one.

^aLecture timestamp: 01:36:09.

3.7 The scale and coupling still missing

Two pieces of bookkeeping have deliberately been postponed. First, X is a length while a_n is dimensionless, so the proportionality in (3.12) must contain a length scale. Second, raising the oscillator level cannot literally add a dimensionless number to a physical mass squared. Restoring the string length ℓ_s , or equivalently the Regge slope $\alpha' \sim \ell_s^2$, supplies the missing dimensions:

$$M^2 \sim \frac{1}{\alpha'}(N - a). \quad (3.40)$$

The precise convention and normalization will be fixed in the next lecture.

Question from the audience. What fixes the unit of string length and the mass added by one creation operator?^a

Response. A dimensionful constant has been suppressed. It is the string scale, usually written ℓ_s or α' . The oscillator operators are dimensionless; powers of that scale convert their level number into physical length, tension, and mass squared.

^aLecture timestamp: 01:37:03.

The second parameter is the string coupling g_s . In the simplest discussion it is one universal number controlling joining and splitting. In richer models, endpoints can carry different labels and strings can meet at junctions. A Y-shaped configuration is the string analogue of three flux tubes meeting; in the old hadronic picture it is associated with a three-quark baryon, whereas a single open segment resembles a quark–antiquark meson.

Question from the audience. Can different string endpoints have different joining probabilities?^a

Response. The simplest model uses one coupling. More elaborate theories attach charges or boundary labels to endpoints, and selection rules then decide which joins are allowed. The detailed construction is deferred, but the universal local joining idea survives.

^aLecture timestamp: 01:39:31.

Question from the audience. Can a string split lengthwise or form a Y-shaped junction?^a

Response. Some string systems do admit junctions. They are not the ordinary two-endpoint splitting process just drawn, but they occur when the theory has compatible string charges. The meson–baryon flux-tube analogy is a useful picture of the distinction.

^aLecture timestamp: 01:40:01.

Question from the audience. Does changing the number of spacetime dimensions change the interaction coupling?^a

Response. Dimensions change the spectrum, phase space, and effective couplings obtained after compactification. The fundamental string coupling still organizes the perturbative expansion, but its lower-dimensional appearance depends on the geometry and volume of the extra dimensions.

^aLecture timestamp: 01:42:25.

Question from the audience. Is gravity what makes a free string move in the first place?^a

Response. No force is required to maintain motion. Initial position and velocity specify a free solution; inertia carries it forward. Gravity or another force changes that motion. Asking what caused the initial conditions is a separate question, not part of the local equation of motion.

^aLecture timestamp: 01:43:17.

We have therefore reached three linked conclusions. Quantizing an open string produces a massless vector at its first excited level; the bosonic ground state below it is unstable; and interactions inevitably introduce closed strings, whose spectrum contains gravity. The next task is to restore the missing scale and perform the closed-string quantization explicitly.

Chapter 4

Closed Strings, Level Matching, and Duality

Overview. A closed string supports independent waves in its two intrinsic directions. Quantizing those waves seems at first to double the open-string photon, but invariance under shifting the origin of the string coordinate removes the one-sided states. The first surviving level contains a massless spin-two particle. The same idea of changing descriptions then opens a broader question: when couplings vary, what does it mean for any object to be fundamental?

4.1 A Generator Kept in Reserve

Before drawing a closed string, let us recover one fact from canonical mechanics. Suppose

$$L = L(q_i, \dot{q}_i), \quad (4.1)$$

and define the momentum conjugate to q_i by

$$p_i = \frac{\partial L}{\partial \dot{q}_i}. \quad (4.2)$$

An infinitesimal continuous transformation can be written

$$\delta q_i = \epsilon f_i(q), \quad (4.3)$$

where ϵ is a small parameter. If this transformation is a symmetry, Noether's theorem supplies a conserved charge. Dividing the coordinate displacement by the common parameter gives the generator

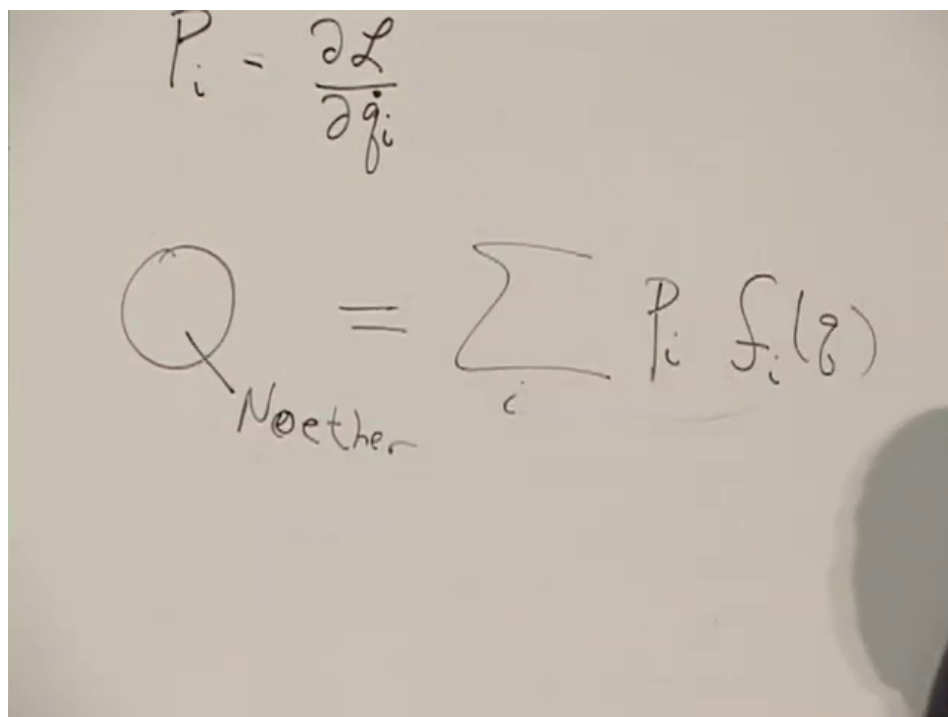
$$Q = \sum_i p_i f_i(q) = \frac{1}{\epsilon} \sum_i p_i \delta q_i. \quad (4.4)$$

In quantum mechanics, Q is the operator that generates the corresponding small transformation.

A rotation in the xy -plane is the familiar example. With one choice of orientation,

$$\delta x = \epsilon y, \quad \delta y = -\epsilon x, \quad (4.5)$$

and (4.4) becomes angular momentum, up to the same choice of sign. The particular example will not matter. What matters is the pattern: momentum contracted with the infinitesimal displacement gives the generator.



The image shows two handwritten equations on a light-colored background. The top equation is $P_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$. The bottom equation is $Q = \sum_i P_i f_i(\tau)$, with a circle around the Q and the word "Noether" written below it.

Figure 4.1: The canonical momentum and Noether generator retained for the later shift of the closed-string coordinate.

Lecture frame: 00:05:00.

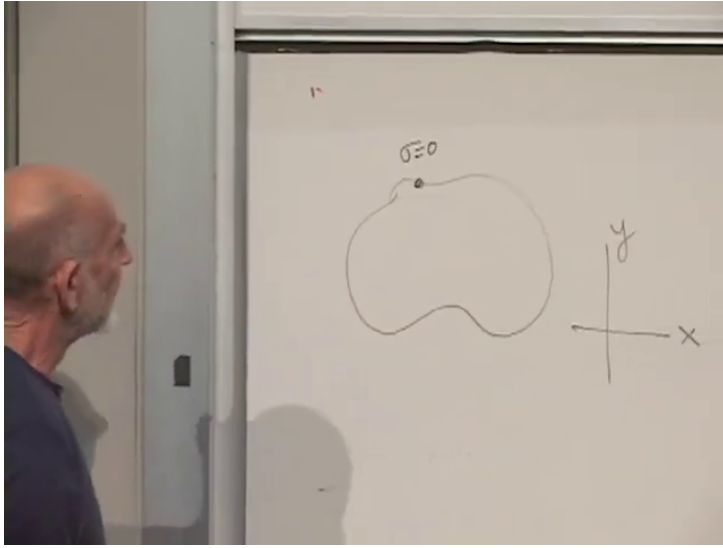
We shall shortly encounter a symmetry whose coordinate is not a point in ordinary space but the labeling of points along a string. Equation (4.4) is being placed within reach for that purpose.

4.2 A Coordinate Around a Loop

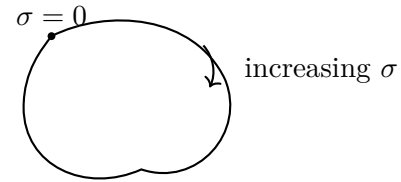
A closed string has no endpoints. Choose a parameter

$$0 \leq \sigma < 2\pi \tag{4.6}$$

and, for the moment, mark one point as $\sigma = 0$. Halfway around is $\sigma = \pi$; a quarter-turn and three-quarter-turn are $\pi/2$ and $3\pi/2$. The mark establishes a coordinate convention, not yet a physical feature.



(a) The marked loop on the blackboard.



(b) The intrinsic orientation.

Figure 4.2: A point is marked to define the origin, and an arrow fixes the direction in which σ increases.

Lecture frame: 00:06:48.

The arrow along the loop is an *intrinsic* orientation. It must not be confused with clockwise or counterclockwise motion in the xy -plane. Imagine a marked rubber band lifted from the board, turned over, and put back. Its apparent clockwise sense reverses, while the sequence of its intrinsic marks does not. Increasing σ remains increasing σ .

This distinction controls the language of right- and left-moving waves. In this chapter, a right-moving disturbance means one traveling toward increasing σ , and a left-moving disturbance means one traveling toward decreasing σ . Neither name tells us which way the disturbance appears to move in the room.

Orientation had a concrete role in the original hadronic string picture. Breaking an oriented flux tube produces a quark at one endpoint and an antiquark at the other; the intrinsic arrow distinguishes the two ends. One can also construct unoriented string theories, but for the present discussion we retain the arrow and the two senses along it.

4.3 Periodicity and Running Waves

Let $X(\sigma, \tau)$ and $Y(\sigma, \tau)$ be two transverse coordinates of the string, with τ denoting time. Extra transverse coordinates, when present, are treated in exactly the same way. Since $\sigma = 0$ and $\sigma = 2\pi$ label the same point,

$$X(\sigma + 2\pi, \tau) = X(\sigma, \tau), \quad Y(\sigma + 2\pi, \tau) = Y(\sigma, \tau). \quad (4.7)$$

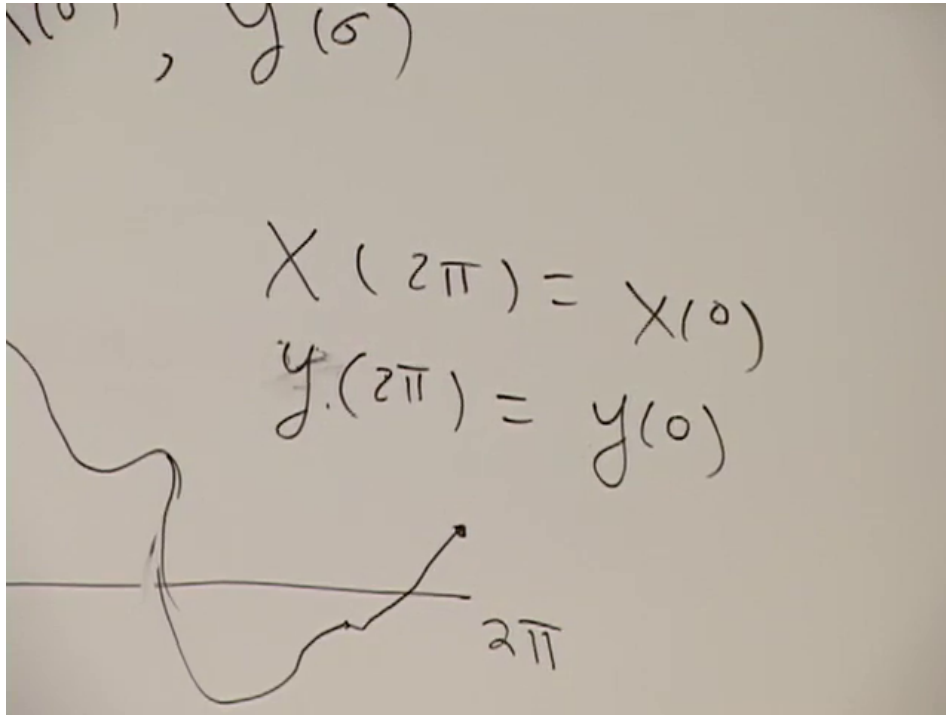


Figure 4.3: Periodic boundary conditions for the two transverse coordinates.

Lecture frame: 00:12:40.

A localized wiggle can therefore circulate indefinitely. When it reaches 2π , it simply reappears at 0. This differs from an open string, where a running wave reaches an endpoint and reflects; what was left-moving becomes right-moving and vice versa. Standing waves are natural for the open string, whereas a closed string admits two independently propagating families.

4.3.1 Complex Fourier modes

The convenient basis for a periodic function is

$$e^{in\sigma} = \cos(n\sigma) + i\sin(n\sigma), \quad n \in \mathbb{Z}. \quad (4.8)$$

The integer condition is forced by periodicity: $e^{in(\sigma+2\pi)} = e^{in\sigma}$. Replacing n by a half-integer, for example, would change the function after one circuit.

Suppressing time dependence momentarily, the blackboard expansion is

$$X(\sigma) = X_0 + \sum_{n>0} X_n e^{in\sigma} + \sum_{n>0} X_{-n} e^{-in\sigma}, \quad (4.9)$$

$$Y(\sigma) = Y_0 + \sum_{n>0} Y_n e^{in\sigma} + \sum_{n>0} Y_{-n} e^{-in\sigma}. \quad (4.10)$$

For real X and Y , the negative-index coefficients are complex conjugates of the positive-index coefficients at the classical level. The zero modes X_0 and Y_0 locate the center of mass; they are positions, not internal oscillations.

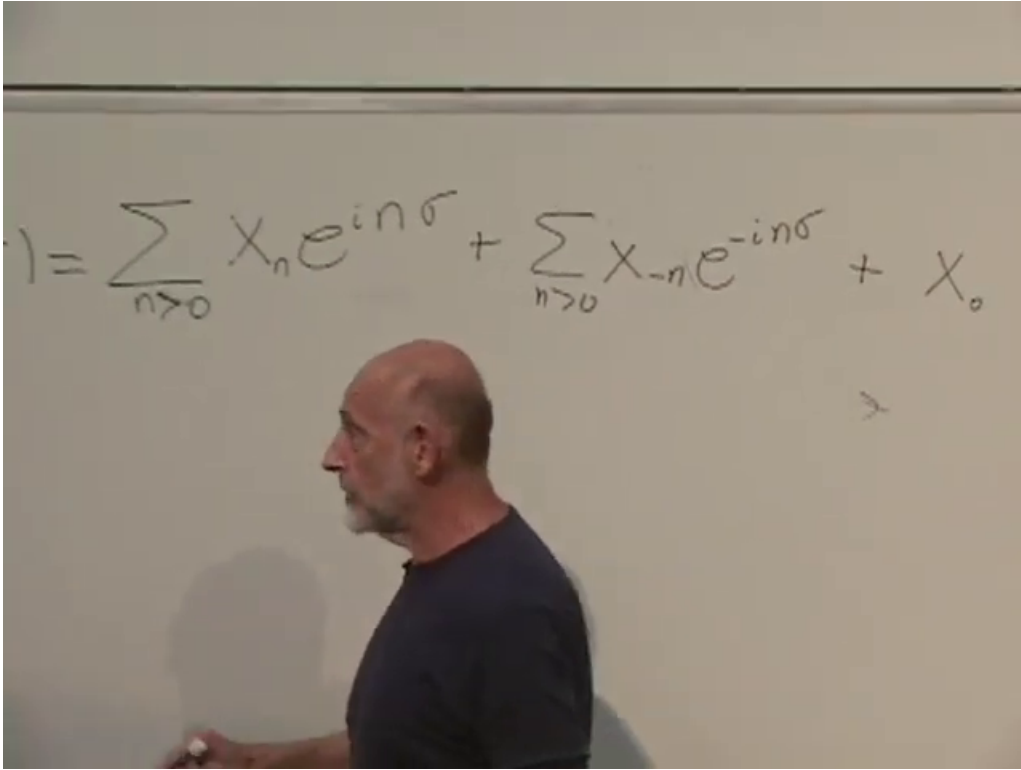


Figure 4.4: The positive- and negative-index pieces of the periodic Fourier expansion, followed by the zero mode.

Lecture frame: 00:20:00.

Question from the audience. If these are moving waves, where is the time coordinate?^a

Response. It has only been suppressed in the board notation. A traveling solution carries a phase involving both σ and τ . With the convention used below, one family has $e^{-in(\tau-\sigma)}$ and the other $e^{-in(\tau+\sigma)}$.

^aLecture timestamp: 00:19:05.

Thus it is the full phase, not the sign of n in a purely spatial Fourier factor by itself, that determines propagation. We may write

$$X(\sigma, \tau) = X_R(\tau - \sigma) + X_L(\tau + \sigma), \quad (4.11)$$

and similarly for Y . Which family is called R rather than L is a convention; the invariant fact is that there are two opposite directions along the σ -circle.¹

4.4 Energy in the Two Directions

The local dynamics is the same as for the open string. In the simplified normalization used here,

$$L \propto \int_0^{2\pi} d\sigma \left[\dot{X}^2 - X'^2 + \dot{Y}^2 - Y'^2 \right], \quad (4.12)$$

¹Editorial clarification: A factor $e^{in\sigma}$ at one instant does not by itself specify a moving wave. Propagation is fixed only after the time phase is supplied.

where a dot denotes ∂_τ and a prime denotes ∂_σ . The energy has a plus sign between kinetic and elastic terms:

$$E \propto \int_0^{2\pi} d\sigma \left[\dot{X}^2 + X'^2 + \dot{Y}^2 + Y'^2 \right]. \quad (4.13)$$

For either transverse coordinate,

$$\dot{X}^2 + X'^2 = \frac{1}{2}(\dot{X} + X')^2 + \frac{1}{2}(\dot{X} - X')^2. \quad (4.14)$$

The cross terms cancel, leaving precisely the original kinetic and gradient energies.

$$E = \int_0^{2\pi} d\sigma \left(\left(\frac{\partial X}{\partial \tau} \right)^2 + \left(\frac{\partial X}{\partial \sigma} \right)^2 \right)$$

$$= \frac{1}{2} \left(\frac{\partial X}{\partial \tau} + \frac{\partial X}{\partial \sigma} \right)^2 + \frac{1}{2} \left(\frac{\partial X}{\partial \tau} - \frac{\partial X}{\partial \sigma} \right)^2$$

Figure 4.5: The total energy reorganized into the two running-wave sectors.

Lecture frame: 00:25:10.

For $X_R(\tau - \sigma)$, one has $\dot{X}_R = -X'_R$; for $X_L(\tau + \sigma)$, one has $\dot{X}_L = X'_L$. Consequently the two squares in (4.14) isolate the two chiral families. Including every transverse coordinate, define

$$E_L \propto \frac{1}{2} \int_0^{2\pi} d\sigma \sum_I (\dot{X}^I + X'^I)^2, \quad (4.15)$$

$$E_R \propto \frac{1}{2} \int_0^{2\pi} d\sigma \sum_I (\dot{X}^I - X'^I)^2. \quad (4.16)$$

Then $E = E_L + E_R$. Because the equations are linear at this stage, waves from the two sectors pass through one another without scattering.

4.4.1 Twice as many oscillators

Substituting the Fourier series into (4.12) turns each nonzero mode coefficient into a harmonic oscillator. Its frequency is proportional to $|n|$: shorter wavelength means higher frequency and therefore a larger excitation energy. Unlike the open string, each positive frequency occurs in two independently propagating copies.

The blackboard labels the X -polarized creation operators a_n^+ and a_{-n}^+ , and the Y -polarized ones b_n^+ and b_{-n}^+ . To separate creation from propagation more cleanly, we shall write

$$a_n^+, \quad b_n^+, \quad \tilde{a}_n^+, \quad \tilde{b}_n^+, \quad n > 0. \quad (4.17)$$

The untilded and tilded operators create opposite world-sheet directions; a and b denote the two transverse polarizations.

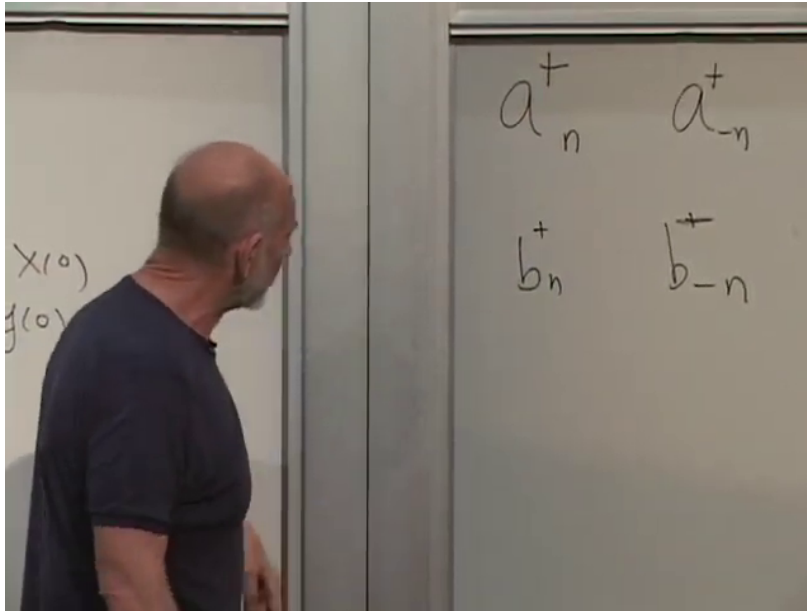


Figure 4.6: The board notation for the doubled X - and Y -polarized creation operators.

Lecture frame: 00:29:10.

If $N_n^{(a)}, N_n^{(b)}$ and their tilded partners are occupation numbers, the oscillator levels carried in the two directions are schematically

$$\begin{aligned} N &= \sum_{n>0} n(N_n^{(a)} + N_n^{(b)} + \cdots), \\ \tilde{N} &= \sum_{n>0} n(\tilde{N}_n^{(a)} + \tilde{N}_n^{(b)} + \cdots). \end{aligned} \quad (4.18)$$

The ellipses stand for additional transverse directions.

4.5 The Spectrum Guess That Fails

Recall the open-string first level. Acting on the ground state with $a_1^\dagger$ or $b_1^\dagger$ gives two linear polarizations. Their circular combinations,

$$c_\pm^\dagger = \frac{1}{\sqrt{2}} (a_1^\dagger \pm ib_1^\dagger), \quad (4.19)$$

carry helicity $h = \pm 1$. The absence of a helicity-zero partner is the signature of a massless vector in this elementary four-dimensional description.

Let the closed-string ground state have mass squared m_0^2 . A first, naive guess is to act once with any one of

$$a_1^\dagger, \quad b_1^\dagger, \quad \tilde{a}_1^\dagger, \quad \tilde{b}_1^\dagger. \quad (4.20)$$

This resembles two copies of the open-string photon, one for each propagation direction. At the next energy, however, the unrestricted list does not assemble into sensible rotational multiplets. The problem is not repaired by more elaborate bookkeeping. The premise that every oscillator monomial is a physical state is wrong.

4.5.1 Level matching, first as a rule

The physical closed-string condition is

$$N = \tilde{N}, \quad \text{equivalently} \quad E_L = E_R. \quad (4.21)$$

For the moment, take this as a postulate; its origin will be derived only after we have seen what it accomplishes.

Every state in (4.20) has one unit of energy in one sector and none in the other, so every one is excluded. The first surviving states carry one oscillator from each family:

$$a_1^\dagger \tilde{a}_1^\dagger |0\rangle, \quad b_1^\dagger \tilde{b}_1^\dagger |0\rangle, \quad (4.22)$$

$$a_1^\dagger \tilde{b}_1^\dagger |0\rangle, \quad b_1^\dagger \tilde{a}_1^\dagger |0\rangle. \quad (4.23)$$

They have total oscillator energy two, divided equally between the two directions. A single $n = 2$ creation operator has the same total oscillator energy but fails (4.21); it contributes two units to only one side.

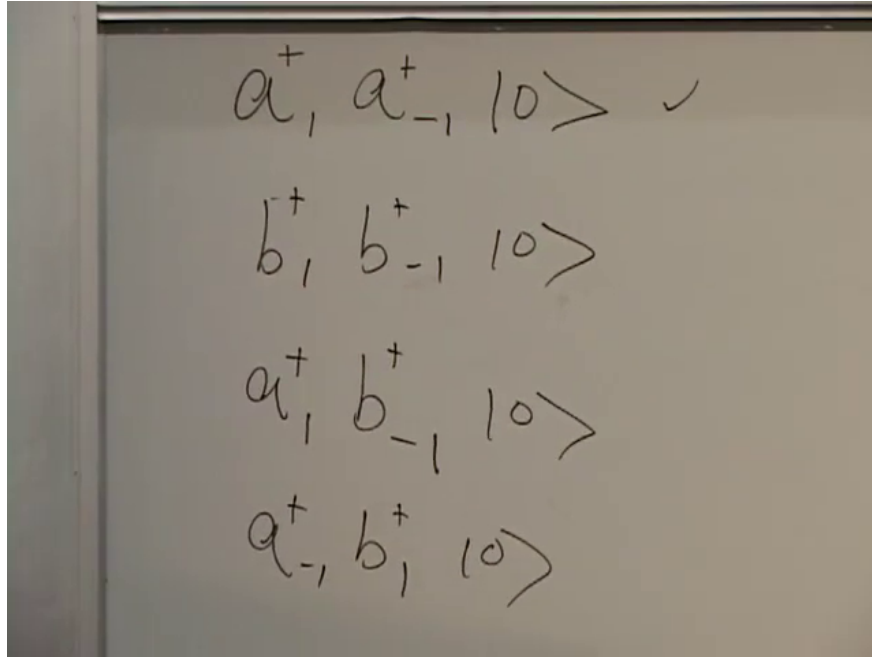


Figure 4.7: The four first-level states that survive level matching. The board uses subscripts 1 and -1 where (4.22)–(4.23) use untilded and tilded operators.

Lecture frame: 00:41:05.

4.6 A Graviton in the First Matched Level

Put both chiral families into a circular basis:

$$c_{\pm}^{\dagger} = \frac{a_1^{\dagger} \pm i b_1^{\dagger}}{\sqrt{2}}, \quad \tilde{c}_{\pm}^{\dagger} = \frac{\tilde{a}_1^{\dagger} \pm i \tilde{b}_1^{\dagger}}{\sqrt{2}}. \quad (4.24)$$

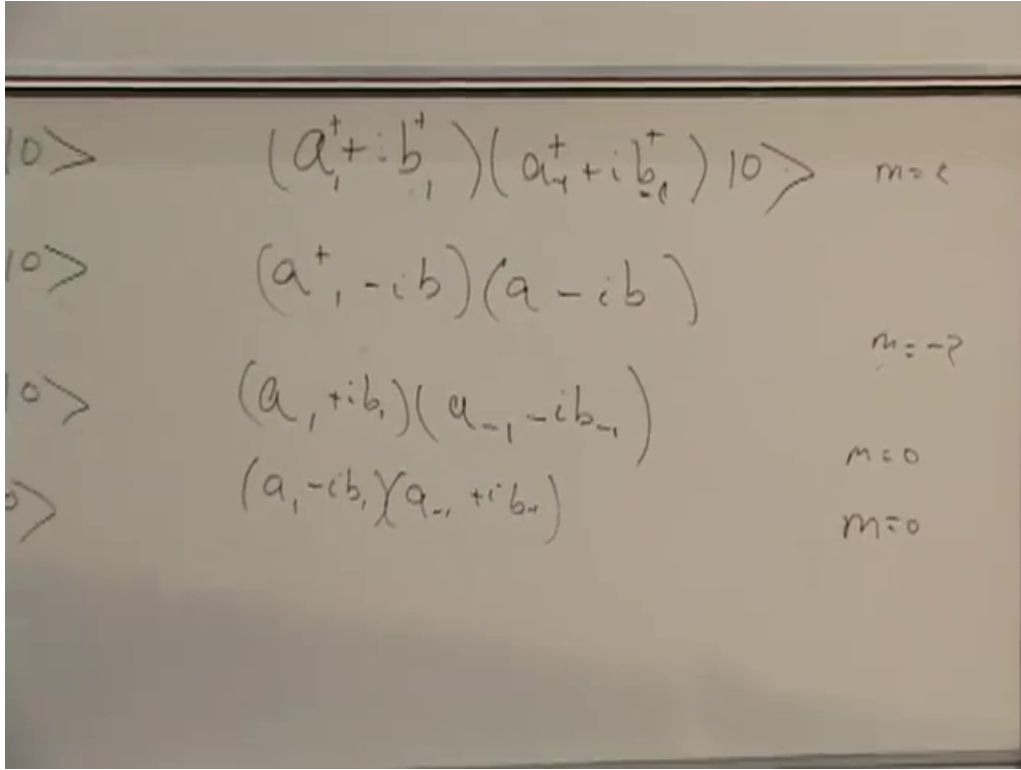
Each factor contributes helicity $+1$ or -1 in ordinary transverse space. Their products therefore give

State	Helicity
$c_+^{\dagger} \tilde{c}_+^{\dagger} 0\rangle$	$+2$
$c_-^{\dagger} \tilde{c}_-^{\dagger} 0\rangle$	-2
$c_+^{\dagger} \tilde{c}_-^{\dagger} 0\rangle$	0
$c_-^{\dagger} \tilde{c}_+^{\dagger} 0\rangle$	0

Question from the audience. Should the second circular-polarization factor come from the opposite propagation sector?^a

Response. Yes. The first expression written on the board was corrected for precisely this reason. A physical state here must pair one right-moving oscillator with one left-moving oscillator; otherwise it is not level matched. The helicities refer to transverse-space polarization, while the two factors refer to opposite world-sheet directions.

^aLecture timestamp: 00:43:20.

Figure 4.8: The four circular-basis products after the board correction, with helicities $+2, -2, 0, 0$.*Lecture frame: 00:46:10.*

A massive spin-two particle has five spin projections,

$$m = 2, 1, 0, -1, -2. \quad (4.25)$$

The matched level contains no $m = \pm 1$ partners. It therefore cannot be a massive spin-two multiplet. A massless spin-two particle, by contrast, has only helicities $+2$ and -2 . The same-helicity products are exactly those two states: they are the graviton polarizations.

Two helicity-zero combinations remain. In the language used in class, one is the dilaton and the other the axion. A covariant description of the generic closed bosonic string refines this statement:

$$\alpha_{-1}^I \tilde{\alpha}_{-1}^J |0; k\rangle \longrightarrow \underbrace{h_{IJ}}_{\text{symmetric traceless}} \oplus \underbrace{B_{IJ}}_{\text{antisymmetric}} \oplus \underbrace{\Phi}_{\text{trace}}. \quad (4.26)$$

These are the graviton, antisymmetric two-form, and dilaton. In four dimensions the two-form can be dualized to an axion-like pseudoscalar, which is the language adopted in the lecture.²

At this formal level the three fields are massless. Compactification, projections, and background dynamics can alter which scalar or two-form degrees of freedom remain light. The graviton is different: once a consistent interacting closed-string sector is present, its massless spin-two state is unavoidable. This is why gravity emerges from closed strings rather than being appended as a separate ingredient.

²Editorial clarification: The second helicity-zero state is fundamentally the antisymmetric tensor field $B_{\mu\nu}$ in the generic closed-string spectrum. Calling it an axion is the four-dimensional dual description.

4.7 Why the Levels Must Match

We now return to the mark at $\sigma = 0$. Is it a physical bead attached to the loop, or merely the origin of our parameter? For an ordinary unmarked closed string, no point should be privileged. The quantum state must be unchanged when every label is shifted together.

4.7.1 The discrete version

Regulate the loop as M ordered points with transverse coordinates (X_i, Y_i) . Temporarily suppress Y_i . A configuration-space wavefunction is

$$\Psi(X_1, X_2, \dots, X_M). \quad (4.27)$$

Changing which point is called first gives a cyclic permutation:

$$\begin{aligned} \Psi(X_1, X_2, \dots, X_M) &= \Psi(X_2, \dots, X_M, X_1) \\ &= \Psi(X_3, \dots, X_M, X_1, X_2). \end{aligned} \quad (4.28)$$

This is not invariance under an arbitrary permutation. Inserting X_3 between X_1 and X_2 , for example, changes the adjacency structure and effectively tears and reconnects the discretized string. Only the common cyclic relabeling

$$X_i \longrightarrow X_{i+1} \quad (\text{mod } M) \quad (4.29)$$

expresses the absence of a preferred origin.

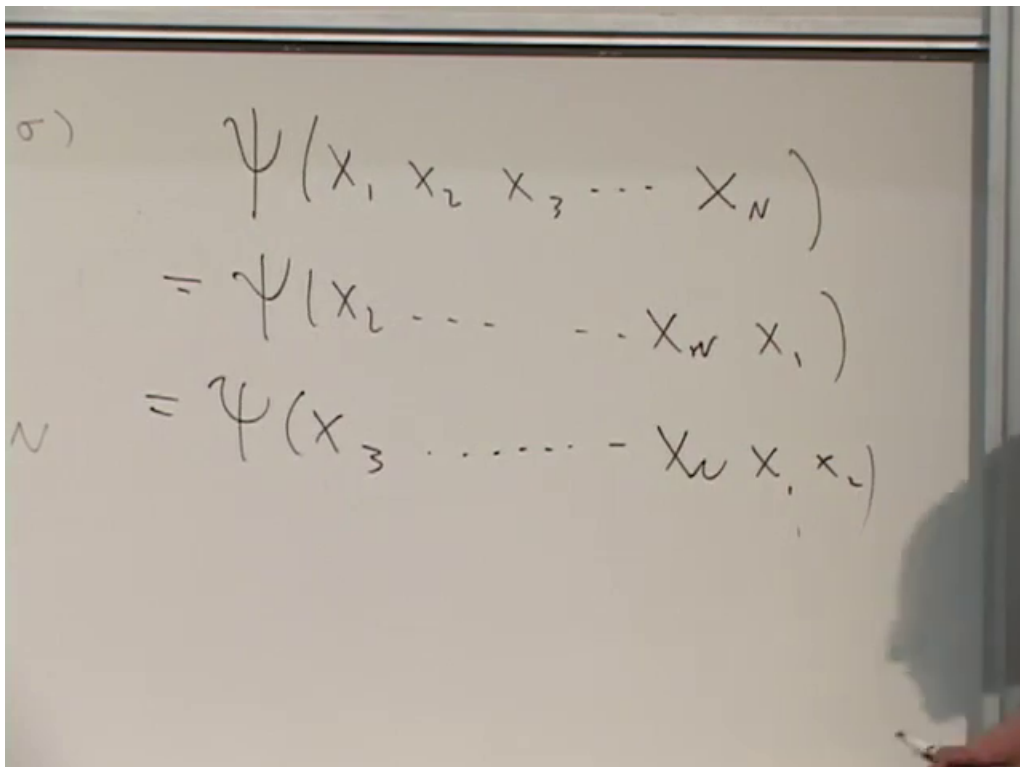


Figure 4.9: The cyclic wavefunction identities for a discretized closed string.

Lecture frame: 00:56:20.

4.7.2 The continuum version

In the continuum limit $M \rightarrow \infty$, the state is the wavefunctional

$$\Psi[X(\sigma), Y(\sigma)].$$

A cyclic relabeling becomes

$$\sigma \longrightarrow \sigma + \epsilon. \quad (4.30)$$

Invariance says

$$\Psi[X(\sigma + \epsilon), Y(\sigma + \epsilon)] = \Psi[X(\sigma), Y(\sigma)]. \quad (4.31)$$

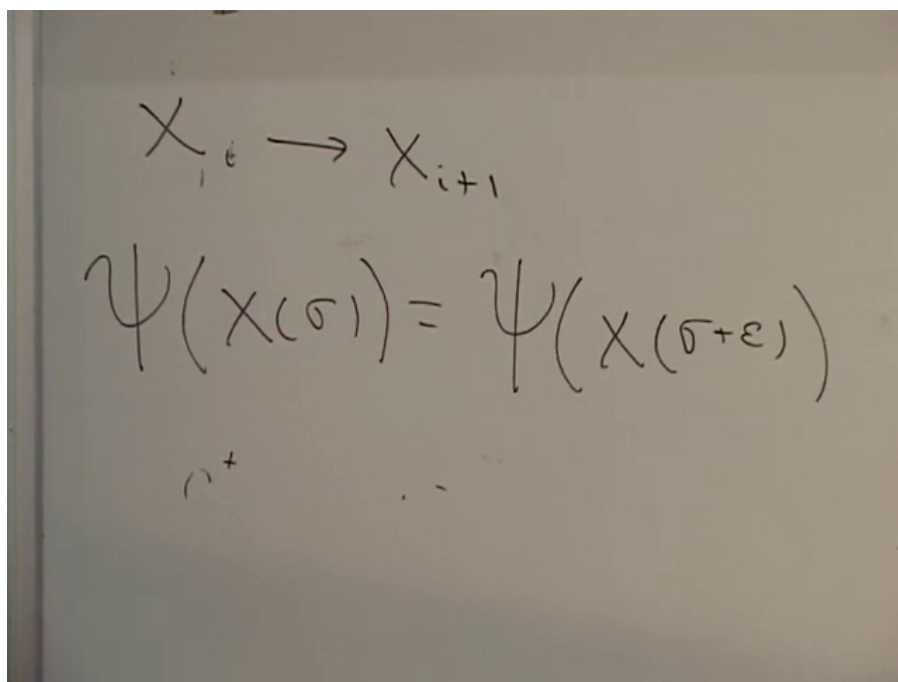


Figure 4.10: The discrete shift and its wavefunctional continuum form.

Lecture frame: 00:58:40.

To first order,

$$\delta X(\sigma) = \epsilon X'(\sigma), \quad \delta Y(\sigma) = \epsilon Y'(\sigma), \quad (4.32)$$

up to the overall sign that distinguishes an active shift of the profile from a passive shift of the coordinate. The corresponding wavefunctional variation is

$$\delta \Psi = \epsilon \int_0^{2\pi} d\sigma \left[\frac{\delta \Psi}{\delta X(\sigma)} X'(\sigma) + \frac{\delta \Psi}{\delta Y(\sigma)} Y'(\sigma) \right]. \quad (4.33)$$

$$\begin{aligned} \circ &= \psi(X(\sigma + \epsilon)) - \\ &\quad - \frac{\partial \psi}{\partial X(\sigma)} \frac{\partial X}{\partial \sigma} \epsilon \end{aligned}$$

Figure 4.11: The first-order expansion of the shifted wavefunctional argument.

Lecture frame: 01:01:31.

Question from the audience. Is the first-order condition a separate equation for each value of σ ?^a

Response. No. The transformation shifts every point of the string together. The variation at one point must be added to the variations at all other points, which is why the continuum expression contains the integral over σ .

^aLecture timestamp: 01:01:30.

In the coordinate representation,

$$P_X(\sigma) = -i \frac{\delta}{\delta X(\sigma)}, \quad P_Y(\sigma) = -i \frac{\delta}{\delta Y(\sigma)} \quad (4.34)$$

with $\hbar = 1$. Equation (4.33) is therefore generated by

$$G_\sigma = \int_0^{2\pi} d\sigma [P_X(\sigma) X'(\sigma) + P_Y(\sigma) Y'(\sigma)]. \quad (4.35)$$

This is exactly the Noether pattern (4.4): canonical momentum multiplied by the coordinate displacement, summed over all degrees of freedom. An unmarked closed-string state obeys

$$G_\sigma |\text{phys}\rangle = 0. \quad (4.36)$$

For the normalization of (4.12), $P_I \propto \dot{X}^I$, so the same condition reads

$$\int_0^{2\pi} d\sigma \sum_I \dot{X}^I X^{I'} = 0. \quad (4.37)$$

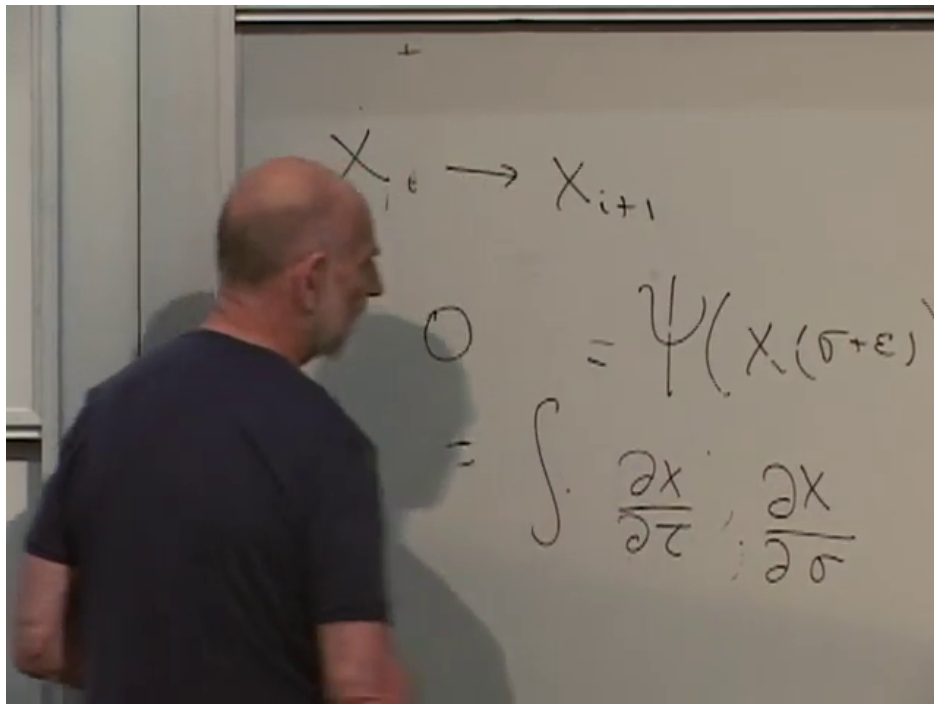


Figure 4.12: The integrated generator after the derivative of the wavefunctional is identified with momentum.

Lecture frame: 01:05:10.

4.7.3 The generator is the energy difference

The sum of the two expressions in (4.16) is the total energy. Their difference is

$$E_L - E_R \propto \frac{1}{2} \int_0^{2\pi} d\sigma \sum_I \left[(\dot{X}^I + X^{I'})^2 - (\dot{X}^I - X^{I'})^2 \right] \quad (4.38)$$

$$= 2 \int_0^{2\pi} d\sigma \sum_I \dot{X}^I X^{I'}. \quad (4.39)$$

The integral is the generator that must annihilate a physical state. Therefore

$$E_L - E_R = 0, \quad \text{or} \quad N - \tilde{N} = 0. \quad (4.40)$$

Level matching is the statement that the origin of the σ -coordinate is unobservable.

Figure 4.13: The energy sum and difference displayed together at the end of the level-matching derivation.

Lecture frame: 01:06:40.

Question from the audience. Why do the displayed formulas appear to place the factor of one-half differently?^a

Response. Only the relative structure is being used. Restoring one consistent overall normalization gives the one-half in the completed-square identity (4.14). The sum yields total energy, while the difference remains proportional to $\int d\sigma \dot{X} X'$. Its vanishing, and hence level matching, is independent of that common convention.

^aLecture timestamp: 01:06:28.

The constraint removes a large portion of the naive oscillator spectrum, and it removes precisely the states that failed to form sensible particle multiplets. It does not remove the open-string photon, because an open string does not possess two independent circulating sectors subject to this closed loop condition. For the closed string, it leaves the graviton, the dilaton, the antisymmetric-tensor or axion-like state, and the higher matched levels.³

4.8 Two Meanings of Left and Right

There are now two unrelated uses of handed language:

³Editorial clarification: In covariant string quantization, level matching is the zero-mode condition $L_0 - \tilde{L}_0 = 0$, one member of the world-sheet constraint system. The derivation here isolates its intuitive origin as invariance under translations around the closed string.

1. $X_R(\tau - \sigma)$ and $X_L(\tau + \sigma)$ describe propagation toward opposite directions along the intrinsic σ -circle.
2. $c_+^\dagger$ and $c_-^\dagger$ describe circular polarization in the physical transverse xy -plane.

The level-matched graviton state contains one oscillator from each world-sheet direction. Those two oscillators may nevertheless have the same spacetime helicity, producing $h = +2$ or $h = -2$. Confusing these two kinds of left and right makes the spectrum look contradictory when it is not.

4.9 What Does Fundamental Mean?

Once open and closed strings have been quantized at this elementary level, a natural question remains: are strings the final constituents, or could they themselves be made of something else?

Question from the audience. Is there evidence or a reason to regard strings as the most fundamental objects, with nothing smaller underneath them?^a

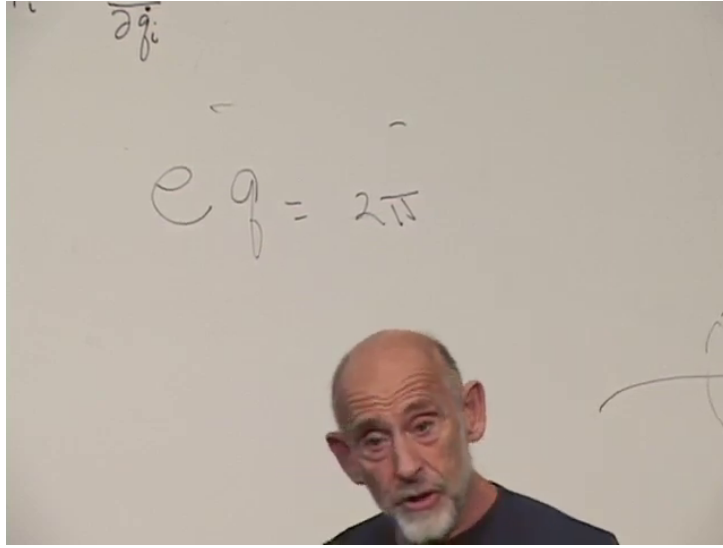
Response. There is no useful parameter-independent answer. The degrees of freedom that look elementary in one regime can become heavy, extended, and apparently composite in another. Fundamental is best read as a statement about which variables provide the useful description at the coupling under consideration.

^aLecture timestamp: 01:10:44.

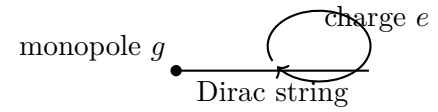
Electric and magnetic charges provide the cleanest model of this answer. Suppose a theory contains an electric charge e and a magnetic charge g . The unobservable Dirac string attached to a monopole requires

$$eg = 2\pi n, \quad n \in \mathbb{Z}. \quad (4.41)$$

For the minimal charges, $n = 1$. A charged particle circling the Dirac string then acquires a phase $e^{ieg} = 1$, so the auxiliary string cannot be detected.



(a) Dirac quantization on the board.



(b) The phase experiment around the string.

Figure 4.14: Electric and magnetic charges are inversely related by the Dirac condition.

Lecture frame: 01:13:20.

If e is small, ordinary perturbation theory in electric variables is useful. Feynman diagrams with additional electric vertices are increasingly suppressed. But (4.41) then makes g large, so a diagrammatic expansion in magnetic vertices is not useful. In terms of the usual dimensionless couplings,

$$\alpha_e = \frac{e^2}{4\pi}, \quad \alpha_m = \frac{g^2}{4\pi}, \quad \alpha_e \alpha_m = \frac{n^2}{4}. \quad (4.42)$$

Weak electric coupling is strong magnetic coupling, and conversely.

The strongly coupled monopole also looks complicated from the electric description. Its field energy is large; it is dressed by an intense cloud of gauge-field fluctuations and virtual magnetically charged pairs. It behaves as a heavy, extended object, a poor starting point for an electric perturbation series. This does not prove that it is less fundamental. It says that it is the wrong elementary variable for that regime.

Now imagine continuously increasing e . The quantization condition forces g to decrease. Near the self-dual region neither description is especially weak, and beyond it the magnetic variables can become the simpler, lighter ones while the electric objects become strongly dressed. Which charge should be called elementary has changed without changing the theory's physical content.

4.9.1 The Wigner question

This ambiguity once surfaced in a discussion of whether the Higgs boson was fundamental or composite. Eugene Wigner repeatedly asked what composite meant. One proposed distinction was decay: perhaps a fundamental object cannot decay and a composite object can. But a ground-state hydrogen atom cannot decay and is plainly composite; the proton may be stable while the neutron decays, yet that does not make one elementary and the other composite in the intended sense. Every simple definition generated another counterexample.

Gerardus 't Hooft ended the discussion with the operational criterion: an object is fundamental when it is useful to think of it as fundamental. This is not an evasion. It recognizes that elementary identifies the variables in which the dynamics is simple, and simplicity can depend on coupling and environment.

4.10 Toward M-Theory

String theory realizes this exchange of roles in a particularly sharp way. In one parameter regime, strings are light and perturbative, while D-branes look like heavy solitonic composites built from string degrees of freedom. As the parameters change, D-branes can become the lighter elementary objects, and configurations described as strings in the first regime can be reinterpreted in terms of them. What looked fundamental can become composite; what looked composite can become fundamental.

The old reductionist picture suggests a one-way descent from large things to smaller constituents and finally to a unique bottom layer. Duality suggests something less linear: descriptions can transform into one another as couplings vary, with no single set of variables privileged everywhere. That is the angle from which the next subject, M-theory, will be approached.

The questions of why the critical dimensions are twenty-six for the bosonic string and ten for the superstring remain open here. Rather than settling every technical issue before moving on, we now change viewpoint and ask how M-theory can produce string theory in an appropriate limit.

Chapter 5

Scales, Strings, and Twenty-Six Dimensions

Overview. What sets the spacing of a string spectrum? The answer is the tension, a quantity with dimensions of mass squared. For fundamental strings that observation leads toward Planck units, the practical impossibility of direct excitation, and the onset of strong gravity. After a deliberate break, we return to the bosonic open-string ground state. Regulating its zero-point energy gives the historical *easy* route to twenty-four transverse directions and hence twenty-six spacetime dimensions.

5.1 The Scale We Had Left Out

So far we have spoken about vibrating and rotating strings without saying what one unit of excitation means. A hadron can be pictured as a string, and its oscillators can be excited, but how large is the jump from one state to the next? What determines it?

There is only one dimensional parameter in the elementary string model. To find it, begin with an ordinary stretched spring. If its endpoints are separated by an amount ℓ , Hooke's law gives

$$E_{\text{spring}} = \frac{k\ell^2}{2}. \quad (5.1)$$

The factor of one-half was supplied by the class after the first version was written. Nothing essential will depend on such numerical factors in this opening estimate.

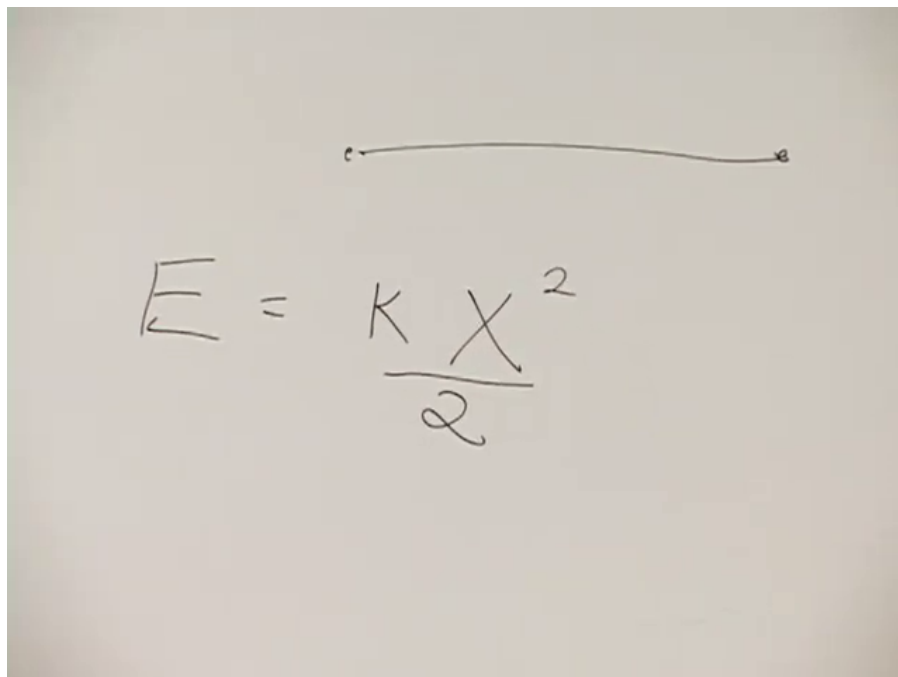


Figure 5.1: The stretched segment and its Hooke-law potential energy.

Lecture frame: 00:02:28.

The energy in (5.1) is nonrelativistic. In the infinite-momentum description used in the preceding lectures, internal nonrelativistic energy is proportional not to m , but to m^2 . We therefore write schematically

$$m^2 \sim k\ell^2, \quad m \sim \sqrt{k}\ell. \quad (5.2)$$

In the string's rest frame its mass is its energy, so the second relation has the form

$$E_{\text{rest}} \sim T_s \ell, \quad T_s \equiv \sqrt{k}. \quad (5.3)$$

The coefficient T_s is the string tension: the energy stored per unit length.

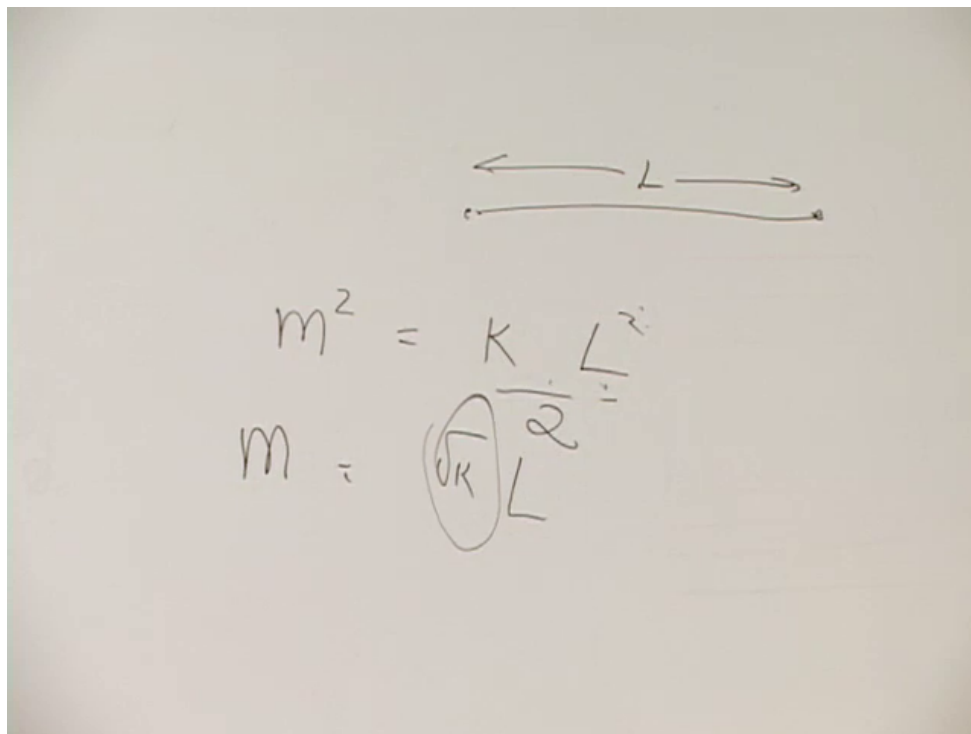


Figure 5.2: The board passes from $m^2 \sim k\ell^2$ to a mass proportional to length.

Lecture frame: 00:05:42.

This notation is adapted to the lecture's spring analogy. In standard relativistic string notation one usually writes

$$T_s = \frac{1}{2\pi\alpha'}, \quad (5.4)$$

so increments of string level produce increments of m^2 of order $1/\alpha'$, equivalently of order the tension.¹

5.1.1 Dimensions and oscillator gaps

Set $c = \hbar = 1$. Energy then has dimensions of inverse length:

$$[E] = L^{-1}. \quad (5.5)$$

Since tension is energy per unit length,

$$[T_s] = L^{-2} = E^2. \quad (5.6)$$

That is why a string spectrum is naturally spaced in mass squared.

¹Editorial clarification: The lecture's effective k is chosen so that $\sqrt{k}$ plays the role of tension. The convention-independent statement is that T_s has dimensions of mass squared and that the Regge slope satisfies $\alpha' \propto T_s^{-1}$.

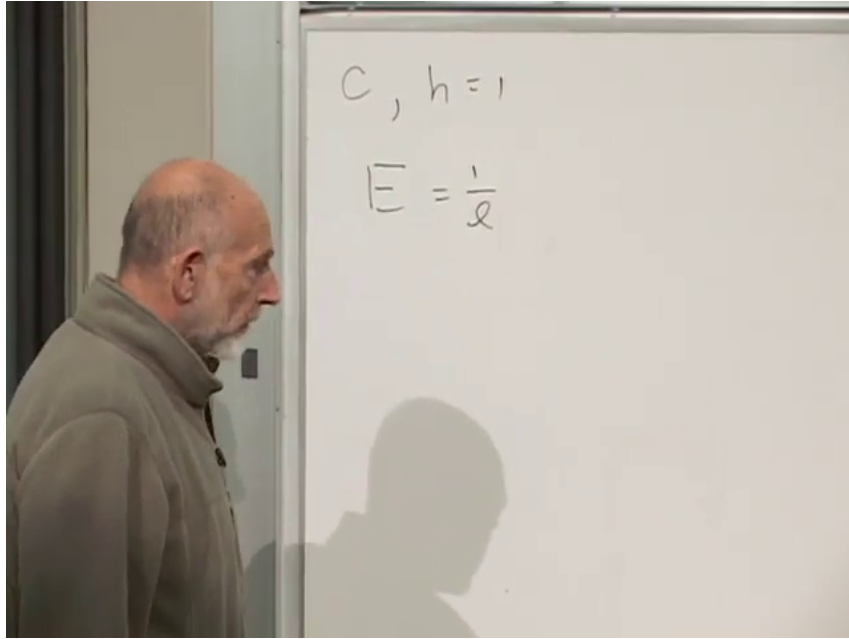


Figure 5.3: Natural units on the board: $c = \hbar = 1$, energy is inverse length, and the oscillator frequency is controlled by the spring scale.

Lecture frame: 00:06:05.

The ordinary oscillator makes the same point from another direction. For a spring with mass μ ,

$$\omega = \sqrt{\frac{k}{\mu}}, \quad \Delta E = \hbar\omega. \quad (5.7)$$

A stiffer spring has a higher frequency and a larger quantum gap. A string contains infinitely many such normal modes, but their common scale is fixed by T_s . Once the tension is known, the spacing of the string spectrum is known.

Tension also has the dimensions of force, because force times distance is energy. Imagine anchoring one endpoint and suspending a weight from the other. An ideal relativistic string with $E = T_s \ell$ can support a force set by T_s , independent of its length. The hadronic tension inferred from observed excitation energies corresponds, in this deliberately fanciful comparison, to the weight of a truck. No truck has ever been hung from a meson; the measured spectrum gives the tension, and the tension gives the mechanical comparison.

Fundamental strings would be far stiffer. If they underlie electrons, photons, and gravitons, their first oscillator excitations must lie far above observed particle energies. In the same fanciful gravitational experiment their tension could support something of galactic mass. The point of the comparison is not engineering. It is that an enormous tension means both a large excitation energy and a very small ground-state fluctuation.

5.2 Units Made from Universal Constants

What scale is enormous enough? The natural benchmark is the Planck scale. Rather than quote it, let us construct it.

Meters, kilograms, and seconds are excellent laboratory units, but their origins are human-sized. A meter is convenient for cloth or rope; a foot is named after a body part; a second is a manageable fraction of a day. One could set the proton radius equal to one and simplify some nuclear formulas, but the proton is only one particular composite particle. Such a choice would not simplify all of physics.

Three dimensional constants have a more universal reach:

- c sets the limiting speed for every material system;
- $\hbar$ appears in the uncertainty relations of every quantum system;
- G couples gravity to every form of mass-energy.

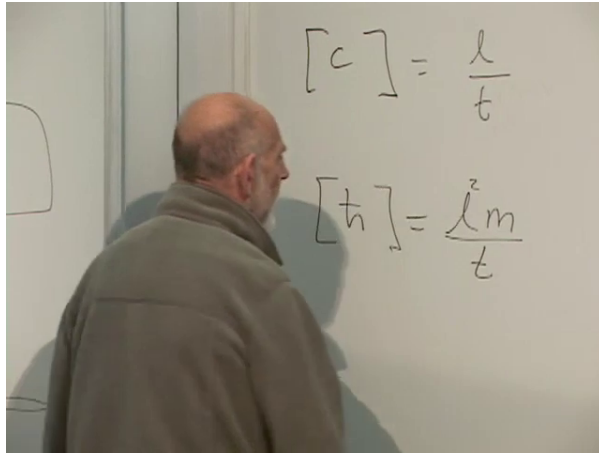
Choosing units in which $c = \hbar = G = 1$ is equivalent to constructing units of length, time, and mass from these constants.

Their ordinary dimensions are

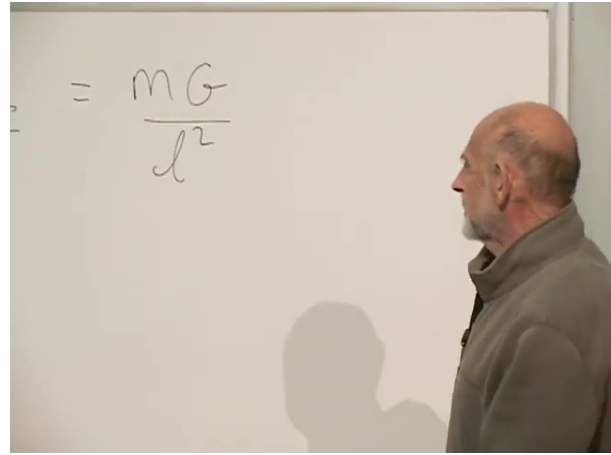
$$\begin{aligned} [c] &= LT^{-1}, & [\hbar] &= ML^2T^{-1}, \\ [G] &= L^3M^{-1}T^{-2}. \end{aligned} \quad (5.8)$$

The middle relation follows directly from $\Delta x \Delta p \gtrsim \hbar$. The last follows, for example, from

$$a = \frac{GM}{r^2}. \quad (5.9)$$



(a) Units of c and $\hbar$.



(b) Units of G from Newtonian acceleration.

Figure 5.4: The dimensional-analysis ingredients assembled in class.

Lecture frame: 00:25:43–00:26:40.

5.2.1 The Planck area first

Seek powers p, q, r for which

$$G^p \hbar^q c^r \quad (5.10)$$

has dimensions L^2 . Choosing area first avoids an early square root and also anticipates the special role that area will later play in gravity. From (5.8), cancellation of mass and time gives

$$-p + q = 0, \quad -2p - q - r = 0. \quad (5.11)$$

The remaining length power must satisfy

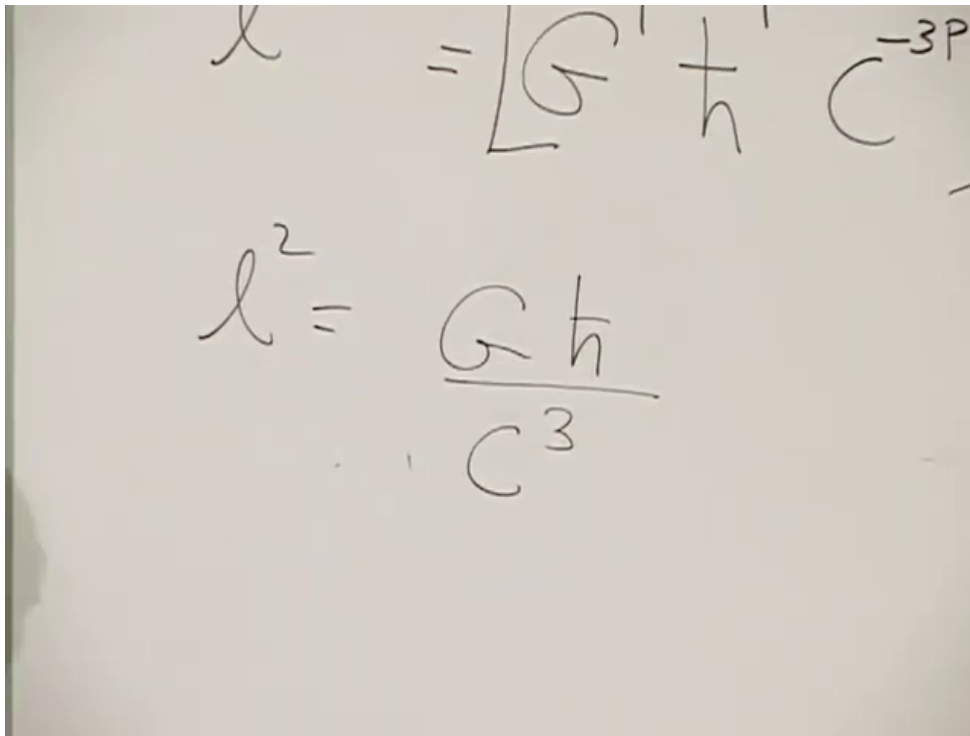
$$3p + 2q + r = 2. \quad (5.12)$$

Thus $q = p$, $r = -3p$, and $p = 1$. The Planck area is

$$\ell_P^2 = \frac{G\hbar}{c^3}, \quad (5.13)$$

and consequently

$$\ell_P = \sqrt{\frac{G\hbar}{c^3}}, \quad t_P = \frac{\ell_P}{c}, \quad m_P = \sqrt{\frac{\hbar c}{G}}. \quad (5.14)$$



The image shows two handwritten equations on a light-colored background. The top equation is $\ell^2 = \frac{G\hbar}{c^3}$, with the variables G , $\hbar$, and c written in a slightly stylized, cursive-like font. The bottom equation is $\ell = \sqrt{\frac{G\hbar}{c^3}}$, also written in a similar style. The equations are centered horizontally and vertically within the image frame.

Figure 5.5: The completed dimensional argument for the Planck area and length.

Lecture frame: 00:31:34.

At order-of-magnitude precision,

$$\begin{aligned} \ell_P &\sim 10^{-35} \text{ m}, & t_P &\sim 10^{-43} \text{ s}, \\ m_P &\sim 10^{-8} \text{ kg}, & m_P c^2 &\sim 10^{19} \text{ GeV}. \end{aligned} \quad (5.15)$$

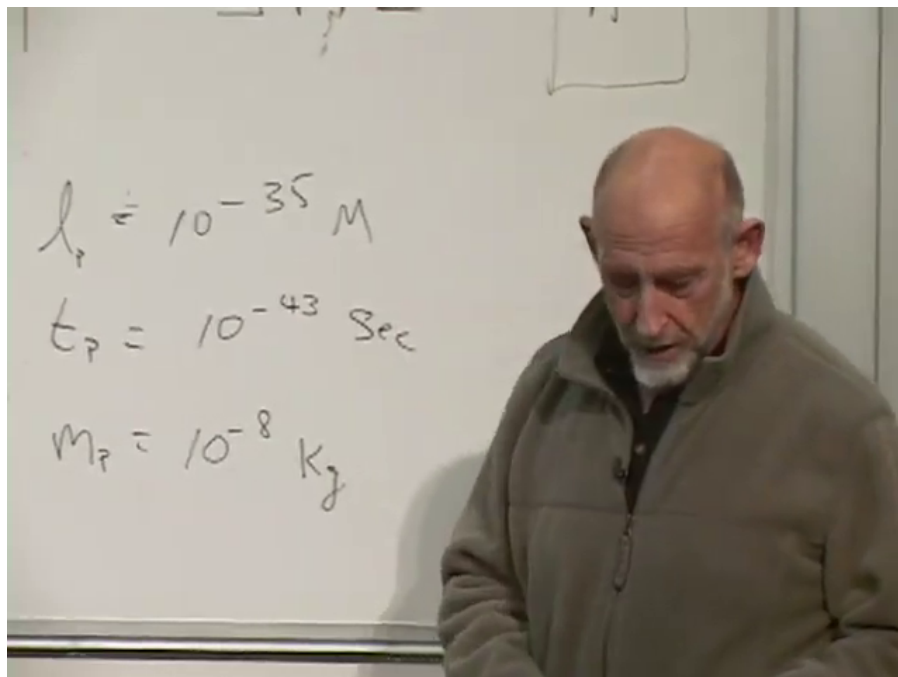


Figure 5.6: The three Planck scales recorded to order of magnitude.

Lecture frame: 00:35:12.

The length and time are unimaginably small, while the mass is surprisingly macroscopic by particle standards, roughly the mass of a fine dust grain. Its rest energy is of order a few gigajoules, comparable to the chemical energy in a tank of gasoline. The observable universe, by contrast, spans roughly 10^{60} Planck lengths and has lived for roughly 10^{60} Planck times. These comparisons are meant to give the units physical weight, not to improve their numerical precision.

If the characteristic string length $\ell_s = \sqrt{\alpha'}$ is near ℓ_P , then the string's quantum size is near ℓ_P , its oscillation period is near t_P , and its first massive excitation is near m_P . The lecture allows factors of ten, one hundred, or more; the identification is an expectation, not a dimensional identity.

Question from the audience. Is the Planck-scale mass the increase above the particle's ground state? ^a

Response. The evenly spaced quantity is mass squared. If the ground state is light, the first string excitation has m^2 larger by an amount of order m_P^2 , so its mass and required center-of-mass energy are of order m_P .

^aLecture timestamp: 00:38:37.

5.3 How Close Should the Two Scales Be?

Question from the audience. Dimensional analysis gives ℓ_P , but why should that number be the length of a string? ^a

Response. It is a physical guess supported by gravity and by the absence of observed string structure, not a consequence of dimensional analysis alone. The string length may be larger than the Planck length, which means its energy scale is lower. The lecture argues only that it should not be an ordinary particle-physics length and that pushing it far below the Planck length runs into gravitational collapse.

^aLecture timestamp: 00:40:21.

The relation between the two scales is in fact coupling dependent. In a complete compactification, ℓ_P , ℓ_s , the string coupling, and the compactification volume enter together. There is no universal equation $\ell_s = \ell_P$. What survives from the lecture is the scale argument: string effects must enter very high, and sub-Planckian localization cannot be discussed while neglecting gravity. ²

Question from the audience. If the Planck length is treated as a smallest useful length, is the Planck mass also the smallest possible mass? ^a

Response. No. Ordinary particles are far lighter. The Planck mass is the scale at which a compact object can have a gravitational radius comparable to its quantum wavelength. In the lecture's heuristic language, it is the scale of the lightest semiclassical black hole, with a lifetime of order a Planck time; the last statement should not be mistaken for a precise description of a fully quantum object.

^aLecture timestamp: 00:43:27.

5.3.1 What running couplings can tell us

There is indirect evidence that conventional quantum field theory remains valid to scales far beyond direct accelerator reach. The three standard-model gauge couplings are measured at accessible energies, and their scale dependence can be extrapolated. In the lecture's schematic plot, the $U(1)$, $SU(2)$, and $SU(3)$ couplings approach one another at a high unification scale, perhaps a few orders of magnitude below the Planck energy.

²Editorial clarification: Modern string constructions need not identify the string length exactly with the four-dimensional Planck length. Their relation depends on the string coupling and compactification data.

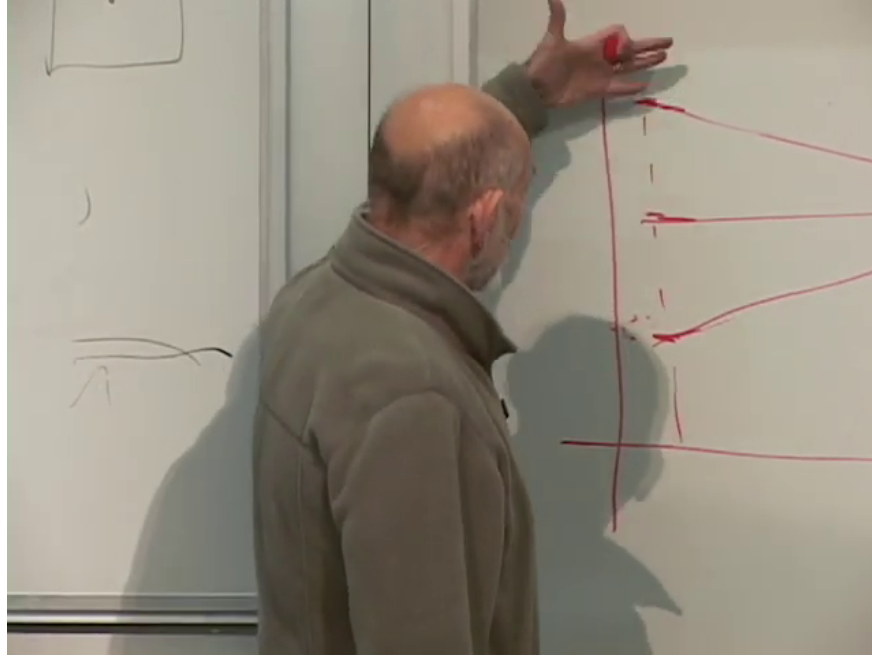


Figure 5.7: The schematic running-coupling plot. The three gauge curves approach one another before the rising gravitational strength becomes order one.

Lecture frame: 00:47:18.

This is an extrapolation, not a direct experiment at the crossing point. Additional particles or thresholds between laboratory and unification energies can change the slopes. The cautious conclusion is only that known field theory shows no sign of string-sized substructure at ordinary distances, and that a plausible string scale lies somewhere in the enormous interval between the inferred unification and Planck scales.

Question from the audience. Which three couplings are being extrapolated? ^a

Response. They are the gauge couplings associated with $U(1)$, $SU(2)$, and $SU(3)$: the hypercharge, weak, and strong interactions. Electromagnetism emerges from electroweak mixing rather than being an independent fourth gauge factor in this accounting.

^aLecture timestamp: 00:48:28.

Question from the audience. Does the gravitational constant G itself increase as we probe higher energy? ^a

Response. G is dimensional, so it is not by itself the correct quantity to compare with a dimensionless gauge coupling. At a characteristic energy M , the dimensionless gravitational strength behaves schematically as $GM^2/(\hbar c)$. That combination grows with energy and becomes order one at the Planck scale.

^aLecture timestamp: 00:49:30.

The force laws make the comparison visible:

$$F_{\text{em}} \sim \frac{e^2}{r^2}, \quad F_{\text{grav}} \sim \frac{GM^2}{r^2}. \quad (5.16)$$

Thus e^2 is compared not with G , but with GM^2 in natural units.

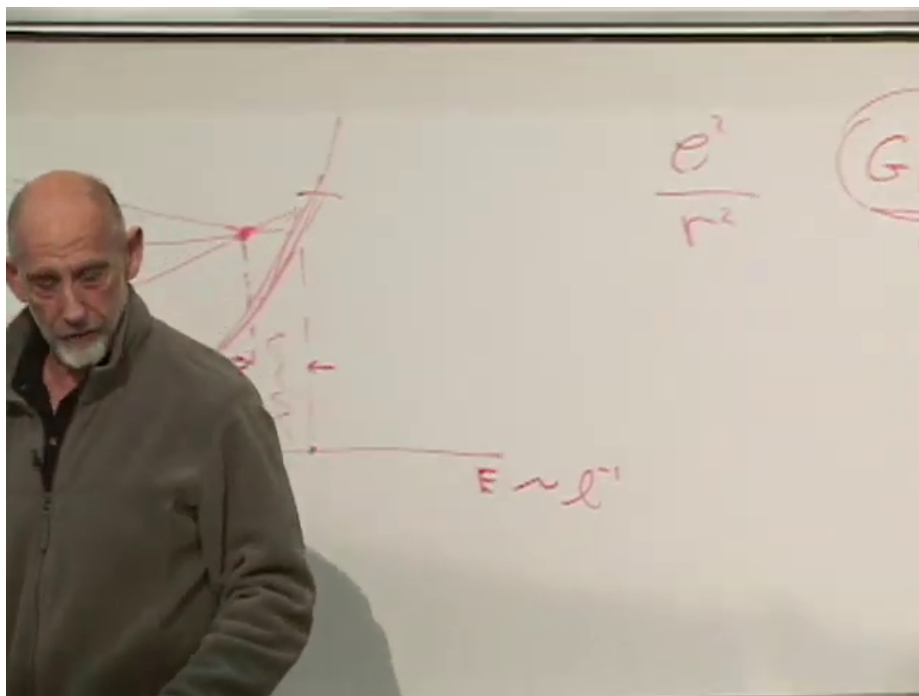


Figure 5.8: Electromagnetic and gravitational force laws used to identify the dimensionless gravitational strength.

Lecture frame: 00:50:50.

5.4 Why Direct String Spectroscopy Is Remote

How would we establish experimentally that an electron is a string? We would do what hadron spectroscopy did: strike it hard enough to excite its internal modes and look for a tower whose masses follow the string pattern. If the first massive level lies near 10^{19} GeV, that is far beyond any direct machine.

The total energy is not the only obstacle. It must be concentrated within a tiny impact parameter. That brings in black holes.

Question from the audience. At these energies would an accelerator begin producing small black holes? ^a

Response. That is the heuristic expectation once center-of-mass energy is concentrated inside its gravitational radius. A stellar black hole is made by gravitational potential energy squeezing a massive star. A microscopic one would instead be made by kinetic energy focused into a tiny collision region. In four-dimensional gravity the required scale is Planckian.

^aLecture timestamp: 00:50:57.

For perspective, a linear accelerator's energy gain is roughly proportional to its length. The lecture takes about two miles for roughly 100 GeV as a simple SLAC-scale yardstick. Reaching 10^{19} GeV at the same gradient would multiply the length by roughly 10^{17} , producing a machine of galactic scale.

Even such a machine would need luminosity. Beams are not aimed with Planck-length precision one pair at a time; enormous numbers of particles cross so that a few collide at the desired impact

parameter. Multiplying a Planck energy by an enormous collision rate produces an equally absurd power requirement. The estimates are intentionally rough, but their conclusion is stable: direct string spectroscopy at the Planck scale is not an experiment one can plausibly design.

Indirect evidence would have to be more structural: a theory that yields the observed low-energy particles and interactions with enough specificity to be tested, or effects of compactification that survive far below the string scale. The lecture does not claim that such a decisive construction is already in hand.

5.4.1 Extra dimensions and the localization limit

Question from the audience. Should the extra dimensions also have sizes near the Planck length? ^a

Response. They are expected to be extremely small in this elementary picture, though not fixed exactly by the argument. If a compact dimension were large enough to explore at lower energy, new Kaluza–Klein modes and higher-dimensional behavior would already have altered the field theory we observe. Its successful extrapolation therefore places an upper bound on such dimensions.

^aLecture timestamp: 00:56:23.

Question from the audience. Why could a string or compact dimension not be many orders of magnitude smaller than ℓ_P ? ^a

Response. Resolving a shorter distance requires a larger momentum. Beyond the Planck scale the energy needed for localization has a gravitational radius comparable to, and then larger than, the region being probed. The experiment makes a black hole before it resolves a smaller ordinary string.

^aLecture timestamp: 00:57:48.

The uncertainty principle gives the estimate. Localizing to Δx requires

Question from the audience. Once Planck length and Planck time are chosen, how is the Planck mass physically determined? ^a

Response. Use either $\hbar$ or G to convert the length scale into an energy. The uncertainty principle says that localizing two particles within ℓ_P requires momentum of order $\hbar/\ell_P$, hence energy of order $m_P c^2$. Equivalently, it is the mass for which quantum and gravitational lengths become comparable.

^aLecture timestamp: 00:59:04.

$$\Delta p \gtrsim \frac{\hbar}{\Delta x}, \quad E \sim c \Delta p. \quad (5.17)$$

At $\Delta x = \ell_P$, this energy is

$$E \sim \frac{\hbar c}{\ell_P} = m_P c^2. \quad (5.18)$$

At the same time the Schwarzschild radius

$$r_s \sim \frac{GE}{c^4} \quad (5.19)$$

is of order ℓ_P , up to numerical constants. Quantum localization and gravitational collapse meet at the same scale.

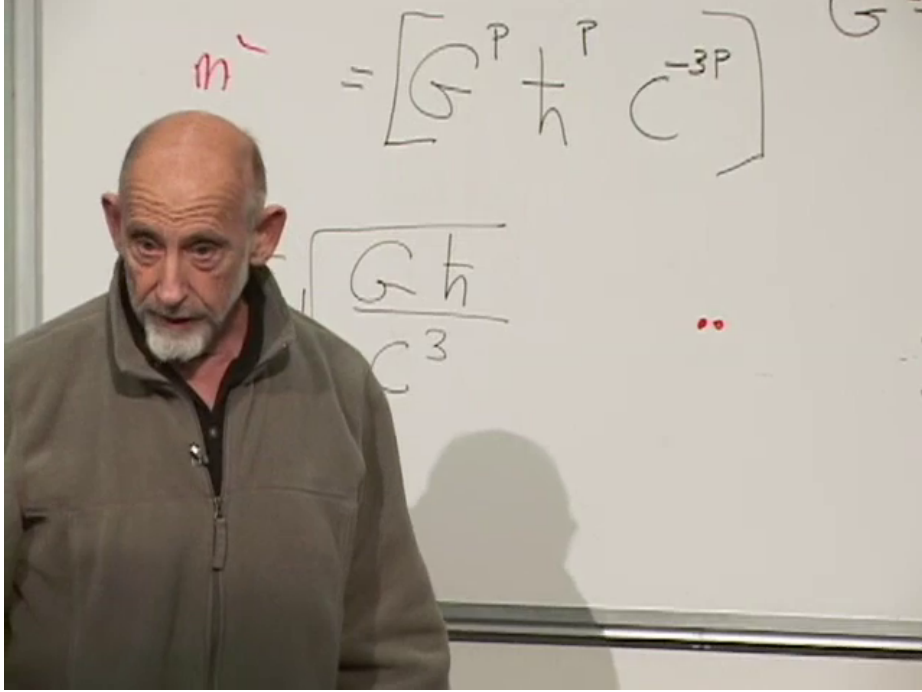


Figure 5.9: The Planck length and mass formulas revisited during the localization argument.

Lecture frame: 01:00:25.

This completes the scale-setting argument.

5.5 The Ground State Must Sit Below the Photon

There are two possible directions after the break: begin conformal transformations, or give the older and easier route to the critical dimension. We choose the latter, with a warning. The easy argument is not easy, and it is not the deepest reason. It is nevertheless correct and close to the historical discovery.

Return to the open bosonic string. Acting once with a transverse creation operator produces two polarization states. Because there is no longitudinal third state, the excitation must be a massless vector. In units where one mode-1 quantum raises m^2 by one,

$$m_{\text{ground}}^2 = -1, \quad m_{\text{first}}^2 = -1 + 1 = 0. \quad (5.20)$$

The tachyonic ground state remains a problem of the bosonic theory; it is acknowledged and set aside. The immediate question is where the negative unit in (5.20) comes from.

Every string mode is a harmonic oscillator. With $\hbar = 1$, its vacuum energy is

$$E_{0,n} = \frac{\omega_n}{2}. \quad (5.21)$$

For the string, $\omega_n = n$, so one transverse coordinate contributes formally

$$E_0^{(1)} = \frac{1}{2} \sum_{n=1}^{\infty} n. \quad (5.22)$$

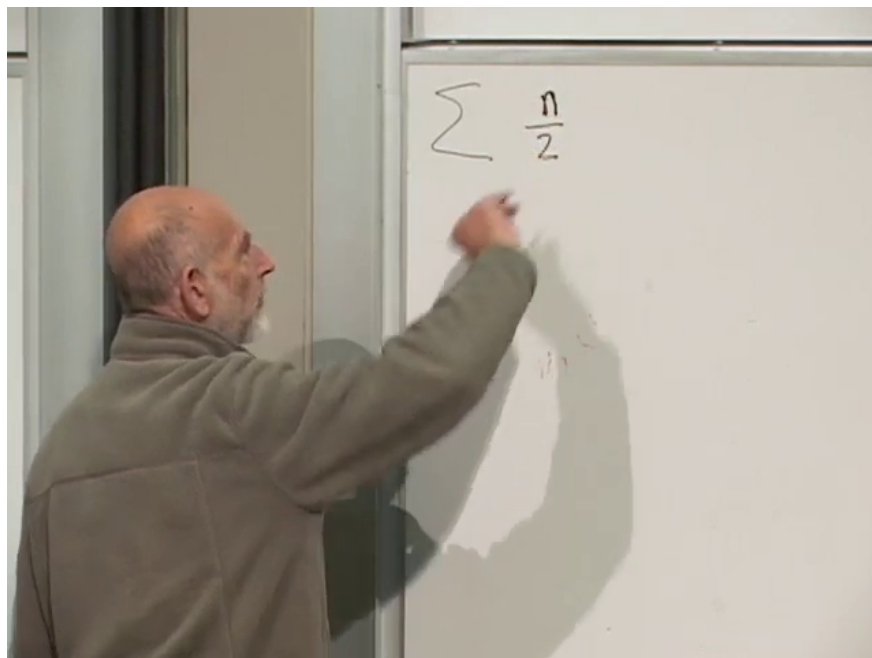


Figure 5.10: The zero-point sum $\sum n/2$ written after the mode frequencies are recalled.

Lecture frame: 01:06:18.

Now the difficulty is sharp. Every term in (5.22) is positive, and the sum diverges. How can it supply a finite negative intercept?

5.6 Subtract the Motion Before Reading the Spectrum

Before regulating the sum, recall why internal energy was identified with mass squared. For a particle carrying a large longitudinal momentum $P > 0$,

$$E = \sqrt{P^2 + m^2} = P + \frac{m^2}{2P} + O(P^{-3}). \quad (5.23)$$

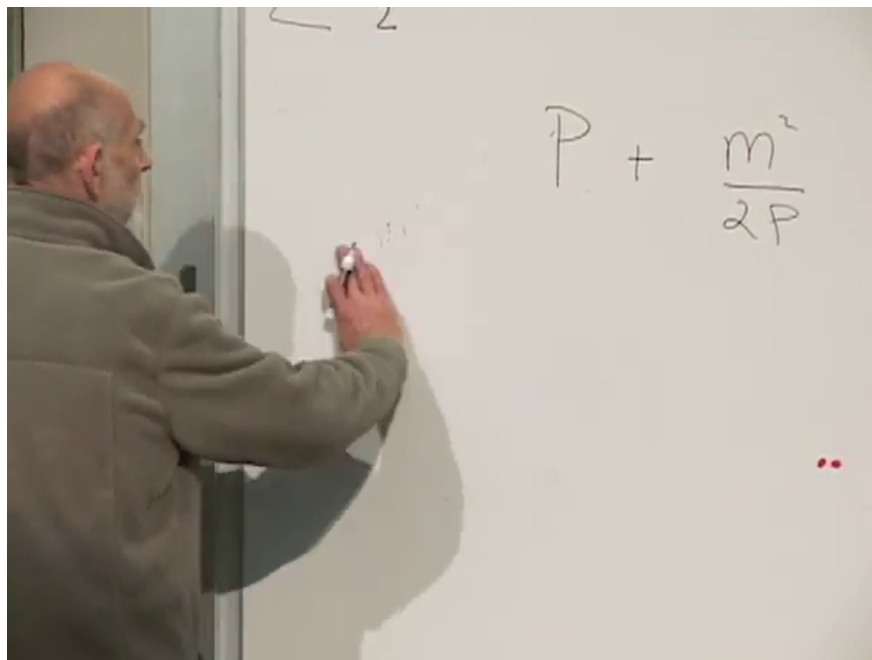


Figure 5.11: The large- P expansion separating overall momentum from the internal mass term.

Lecture frame: 01:07:35.

The first term is the uninteresting overall motion down the z -axis. It is conserved and common to all internal states with the same P . Subtracting it and rescaling removes the kinematic time dilation:

$$\begin{aligned} E - P &= \frac{m^2}{2P} + O(P^{-3}), \\ 2P(E - P) &= m^2 + O(P^{-2}). \end{aligned} \tag{5.24}$$

The atom provides the physical picture. Boost a hydrogen atom. Its total energy grows, but the energy difference between two internal levels, as measured in the laboratory time, shrinks because the atom's clock is dilated. The factor $1/P$ in (5.23) is that suppression. Multiplying by $2P$ restores the invariant mass-squared difference.

The raw ground-state energy therefore need not be literally -1 . It may contain a large state-independent piece associated with the overall longitudinal energy, plus a finite internal remainder. The lecture rewrites the boosted relation to display both pieces before turning back to the oscillator sum.

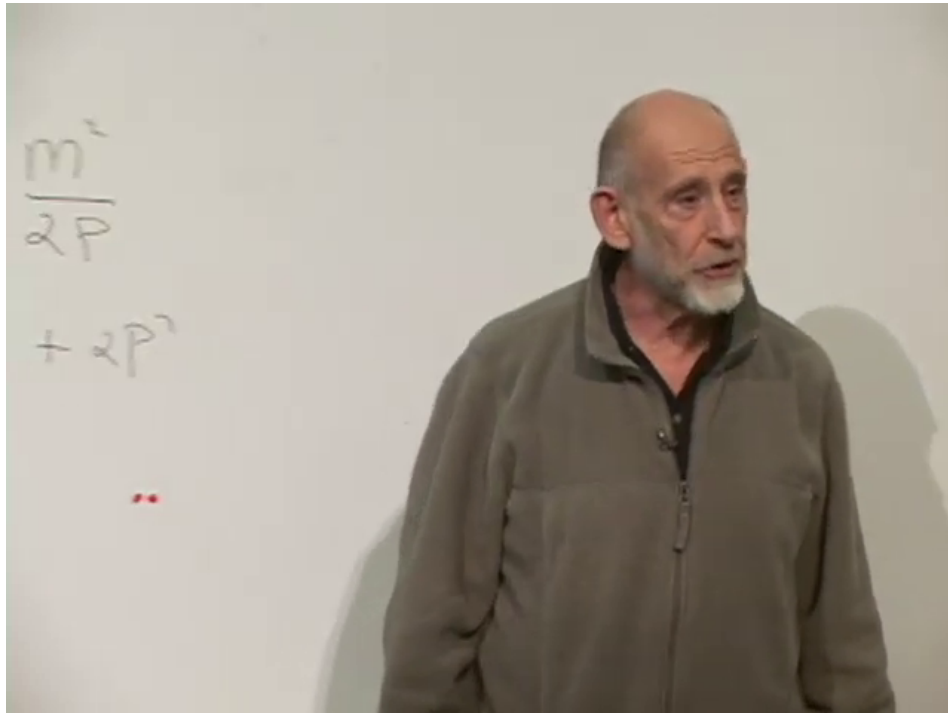


Figure 5.12: The invariant mass term displayed beside the large state-independent longitudinal contribution.

Lecture frame: 01:11:34.

Infinite vacuum energies are familiar in quantum field theory. One often separates a divergent state-independent constant from observable energy differences. The lecture compares this with the zero-point energy of the electromagnetic field in the room and openly calls the procedure unsatisfactory but standard. Here we follow the same heuristic: isolate the divergent part and ask whether a definite finite remainder survives.

5.7 Regulating the Sum

Introduce a positive regulator ϵ :

$$S_1(\epsilon) = \sum_{n=1}^{\infty} n e^{-n\epsilon}. \quad (5.25)$$

For every $\epsilon > 0$, the exponential eventually defeats the factor n , so the sum converges. Sending ϵ toward zero removes the cutoff.

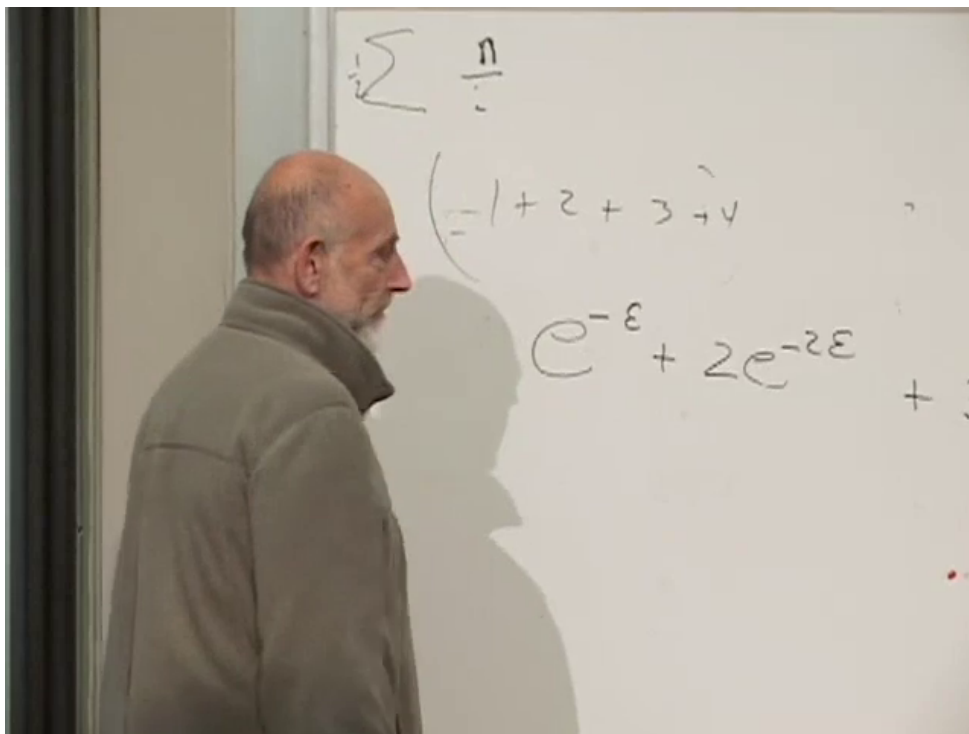
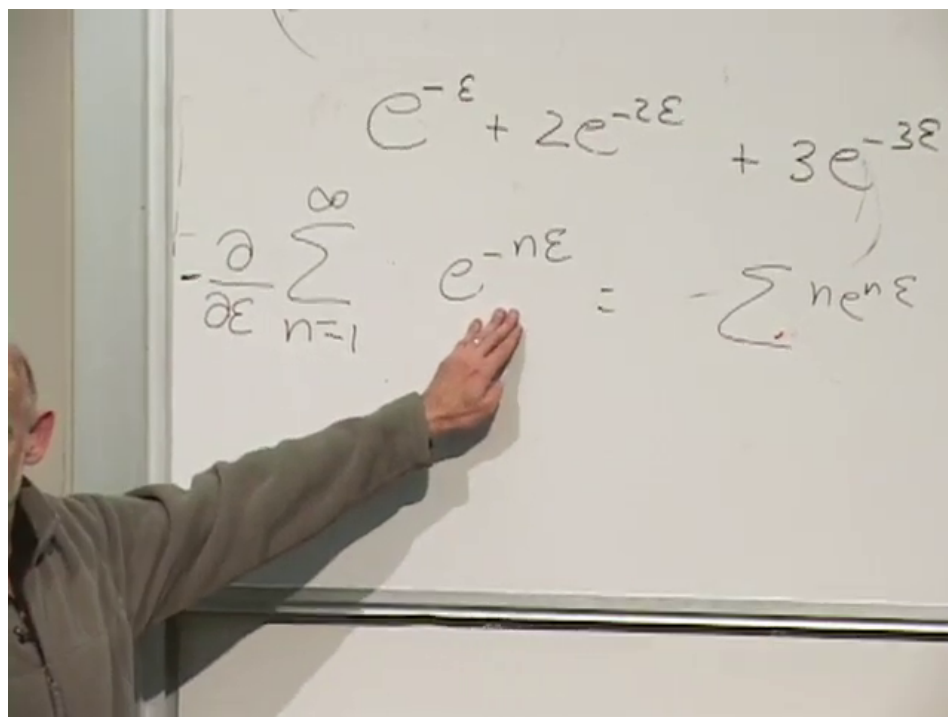


Figure 5.13: The exponential regulator applied term by term to $1 + 2 + 3 + \dots$.

Lecture frame: 01:16:42.

The factor n can be generated by differentiation:

$$S_1(\epsilon) = -\frac{d}{d\epsilon} \sum_{n=1}^{\infty} e^{-n\epsilon}. \quad (5.26)$$

Figure 5.14: Differentiation with respect to the cutoff brings down $-n$.*Lecture frame: 01:18:05.*

The remaining series is geometric:

$$S_0(\epsilon) \equiv \sum_{n=1}^{\infty} e^{-n\epsilon} = \frac{e^{-\epsilon}}{1 - e^{-\epsilon}} = \frac{1}{e^{\epsilon} - 1}. \quad (5.27)$$

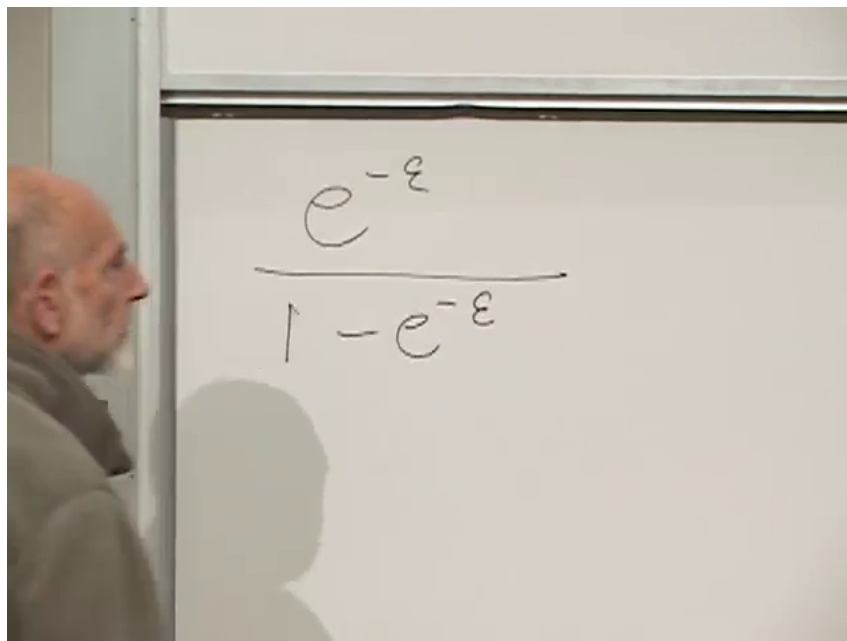


Figure 5.15: The regulated geometric series in closed form.

Lecture frame: 01:19:12.

Expand at small ϵ :

$$S_0(\epsilon) = \frac{1}{\epsilon} - \frac{1}{2} + \frac{\epsilon}{12} - \frac{\epsilon^3}{720} + O(\epsilon^5). \quad (5.28)$$

The board obtains this by expanding numerator and denominator separately, factoring one power of ϵ , and then expanding the reciprocal. That calculation contains several in-class sign checks; the compact expression (5.28) is its cleaned result.

$$\frac{1}{\epsilon} \quad \frac{\left(1 - \epsilon + \frac{\epsilon^2}{2} \dots\right)}{\epsilon - \frac{\epsilon^2}{2} + \frac{\epsilon^3}{6}}$$

Figure 5.16: The numerator and denominator expansions before the reciprocal is simplified.

Lecture frame: 01:23:10.

Applying $-d/d\epsilon$ gives

$$S_1(\epsilon) = \frac{1}{\epsilon^2} - \frac{1}{12} + \frac{\epsilon^2}{240} + O(\epsilon^4). \quad (5.29)$$

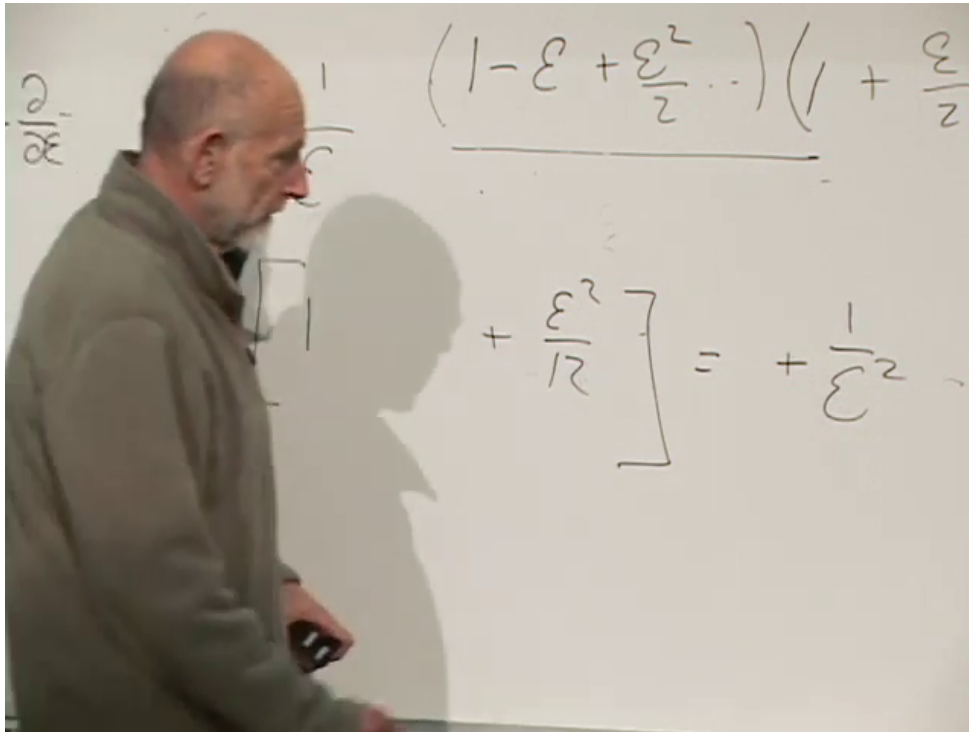


Figure 5.17: The divergent $1/\epsilon^2$ term and finite $-1/12$ remainder at the end of the regulator calculation.

Lecture frame: 01:29:48.

It is common to summarize this result jokingly as $1 + 2 + 3 + \dots = -1/12$. That is not ordinary convergence. Equation (5.29) says that this particular regulator produces a divergent term plus a regulator-independent finite part:

$$\underbrace{\frac{1}{\epsilon^2}}_{\text{divergent}} + \underbrace{\left(-\frac{1}{12}\right)}_{\text{finite}} + O(\epsilon^2). \quad (5.30)$$

The lecture absorbs the first term into the already-present state-independent piece and retains the second. It also preserves the warning attributed to Pauli: being infinite does not by itself make a quantity zero. A consistent theory must explain why the subtraction is allowed.

Restoring the factor of one-half in (5.22), one transverse coordinate contributes

$$E_0^{(1)} = -\frac{1}{24}. \quad (5.31)$$

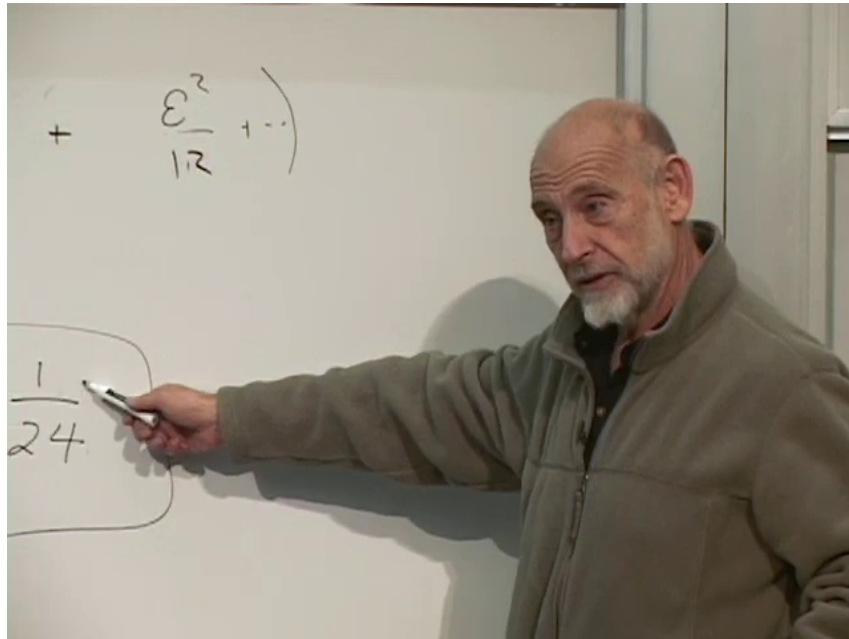


Figure 5.18: The factor of one-half converts the finite $-1/12$ into $-1/24$ per transverse coordinate.

Lecture frame: 01:33:42.

5.8 Counting the Transverse Directions

Let $D_{\perp}$ be the number of transverse string coordinates. Their total vacuum contribution is

$$E_0^{\text{open}} = -\frac{D_{\perp}}{24}. \quad (5.32)$$

For the first open-string oscillator to raise the ground state from -1 to the massless level, we need

$$-\frac{D_{\perp}}{24} = -1. \quad (5.33)$$

Therefore

$$D_{\perp} = 24. \quad (5.34)$$

Those are the directions perpendicular to the large longitudinal momentum. Adding that longitudinal spatial direction and time gives

$$D = D_{\perp} + 2 = 26. \quad (5.35)$$

Question from the audience. When did this argument enter string theory? ^a

Response. It belongs to the original hadronic-string era, around 1969 and the early 1970s. Regge trajectories were already known experimentally. Once the oscillator model was explored on its own, the first excited state was seen to have two rather than three vector polarizations, forcing it to be massless. The required intercept then led quickly to twenty-four transverse oscillators.

^aLecture timestamp: 01:35:27.

At the time this looked almost too strange to believe. Why should a regularized vacuum sum know the dimension of spacetime? Yet the pieces fit: the massless vector, the intercept, and the transverse oscillator count all selected the same number.

5.8.1 The closed string checks the count

A closed string needs one left-moving and one right-moving mode to reach its first physical massless level. Its ground state must therefore begin two units below zero.

Question from the audience. If twenty-four transverse directions give only -1 , must a closed string live in forty-eight transverse directions? ^a

Response. No. The same spacetime contains both open and closed strings. A closed string already has twice the oscillator content: one tower circulating in each world-sheet direction. Each tower contributes $-24/24 = -1$, so together they give -2 .

^aLecture timestamp: 01:37:47.

Thus

$$E_0^{\text{closed}} = 2 \left(-\frac{D_{\perp}}{24} \right) = -2 \quad (D_{\perp} = 24). \quad (5.36)$$

Adding one oscillator from each sector raises the state by two units and lands at zero mass squared, exactly where the graviton multiplet appeared in the preceding lecture.

This agreement makes the argument smell better, but it is still not the deep explanation. In modern quantization the normal-ordering constant, closure of the Lorentz or Virasoro algebra, and cancellation of the conformal anomaly belong to one consistent structure. That structure fixes both the intercept and the critical dimension; the finite cutoff remainder is the historical way into it. ³

That deeper route is next. Only a little complex analysis is needed to begin: complex coordinates, analytic functions, and the Cauchy–Riemann equations. The promised physical image is Turkish taffy—pulled one way, stretched another, reshaped without destroying the structure that matters. The mathematics behind that image is conformal invariance.

³Editorial clarification: The regulator calculation is a heuristic view of the normal-ordering constant. The deeper derivation of $D = 26$ requires world-sheet conformal invariance, equivalently cancellation of the conformal anomaly and consistent constraint algebra.

Chapter 6

Scattering, Duality, and the Veneziano Amplitude

Overview. Scattering reduces a collision to invariant data that can be compared with experiment. For four identical particles those data may be organized into the Mandelstam variables s, t, u , only two of which are independent. Ordinary particle exchange then suggests separate direct and crossed channels. The Veneziano amplitude overturned that expectation: one function contains both channel descriptions. The second half of the lecture builds that function from two open strings that join, evolve as one oscillating chain, and split again.

6.1 What a Theory Must Predict

Fermionic strings solve two defects of the purely bosonic model: they admit fermionic states and, after the appropriate projection, remove the tachyon. Their critical spacetime dimension is ten rather than twenty-six. Extra dimensions therefore remain, but their compactification is postponed. The immediate subject is scattering.¹

Scattering is not necessarily the most interesting thing particles do. It is simply the operation that an experiment can repeat cleanly. Send specified particles toward one another, record the particles emerging from the collision, and infer the dynamics hidden inside the interaction region. A theory becomes quantitative when it supplies a complex amplitude

$$\mathcal{A}(\text{incoming data} \rightarrow \text{outgoing data}), \quad (6.1)$$

whose squared magnitude, together with the usual phase-space factors, gives the probability for that outcome.

¹Editorial clarification: The tachyon-free statement refers to the consistent superstring spectrum after the required projection; merely adding world-sheet fermions is not by itself enough.

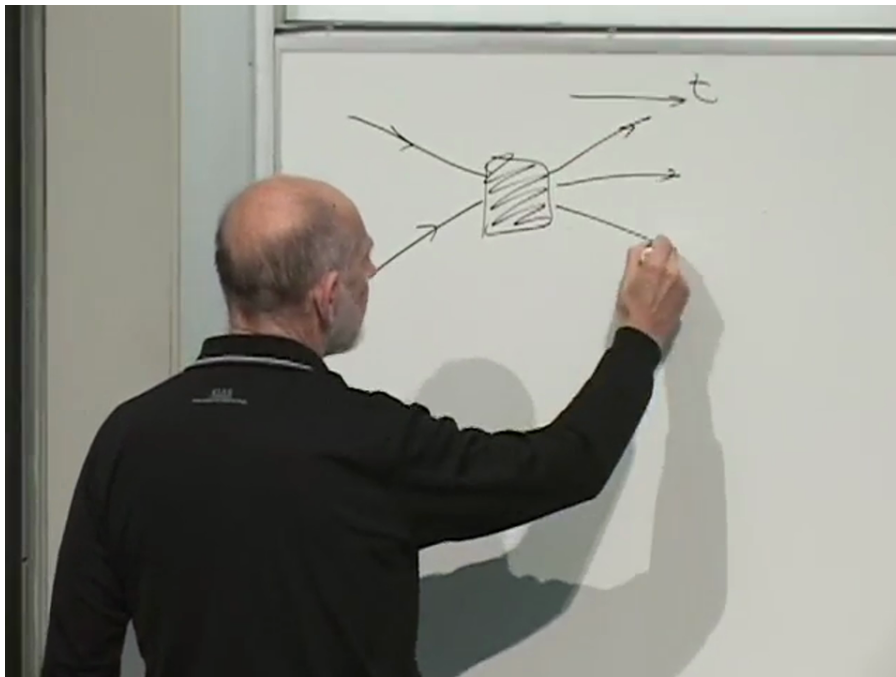


Figure 6.1: Incoming particles enter a black-box interaction region and outgoing particles emerge; time is drawn horizontally.

Lecture frame: 00:02:35.

Particles also carry spin, charge, and species labels. To expose the kinematics first, let us temporarily suppress everything except momentum.

6.2 Four-Momentum and the All-Incoming Convention

Write the four-momentum of one particle as

$$k^\mu = (k^0, \mathbf{k}) = (E, \mathbf{p}). \quad (6.2)$$

The lecture uses the mostly-plus metric, so

$$k^2 \equiv \mathbf{p}^2 - E^2. \quad (6.3)$$

The relativistic dispersion relation is therefore

$$E^2 - \mathbf{p}^2 = m^2, \quad k^2 = -m^2. \quad (6.4)$$

Changing the spatial momentum changes the energy, but it does not move the particle off its mass shell.

$$\begin{aligned}
 E &= p + m \\
 E^2 - p^2 &= m^2 \\
 p^2 - E^2 &= -m^2 \\
 p_\mu p^\mu &= k^2
 \end{aligned}$$

Figure 6.2: The mass-shell condition in the metric convention used throughout the lecture.

Lecture frame: 00:05:56.

Consider two particles entering with momenta k_1, k_2 and two leaving with physical momenta q_3, q_4 . Conservation of energy and momentum is one four-vector equation,

$$k_1 + k_2 = q_3 + q_4. \quad (6.5)$$

It is convenient to assign an incoming momentum to every external line:

$$k_3 \equiv -q_3, \quad k_4 \equiv -q_4. \quad (6.6)$$

Then

$$k_1 + k_2 + k_3 + k_4 = 0, \quad (6.7)$$

and every leg obeys $k_i^2 = -m_i^2$. The outgoing particles have negative energy only in this bookkeeping convention; their physical energies remain positive.

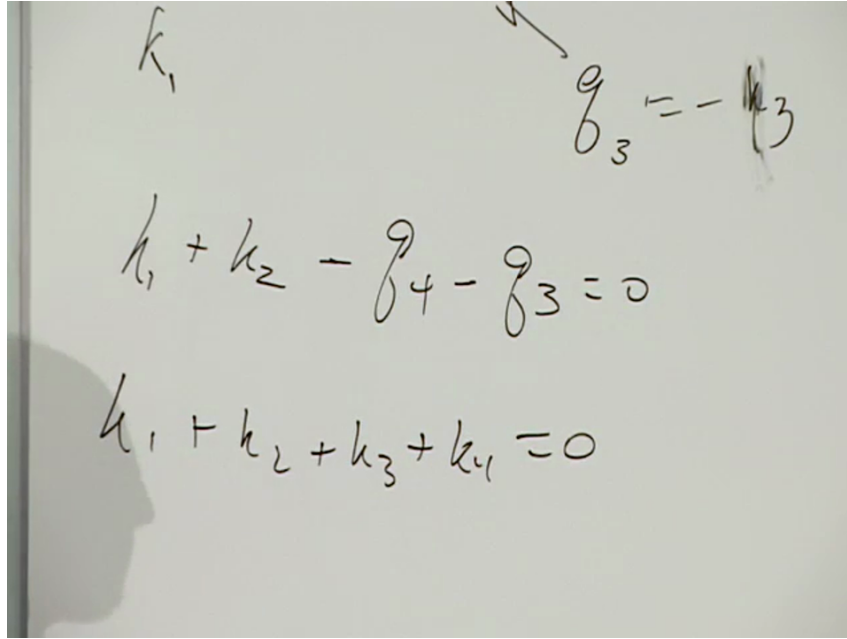


Figure 6.3: Physical outgoing momenta are reversed so that all four external momenta may be treated symmetrically as incoming.

Lecture frame: 00:09:10.

At first the amplitude appears to be a function of sixteen components, $\mathcal{A}(k_1, k_2, k_3, k_4)$. Most of them are redundant. Go to the center-of-mass frame, rotate the incoming momenta onto a chosen axis, and assume for simplicity that all four particles have the same mass. The incoming state is then fixed by one number, the center-of-mass energy. The outgoing spatial momenta must be equal and opposite and have the same magnitude as the incoming pair. Only their deflection angle θ remains. The process therefore has two independent kinematic variables:

$$E_{\text{cm}}, \quad \theta. \quad (6.8)$$

6.3 Mandelstam Variables

The two physical quantities in (6.8) should be expressible without choosing a frame. With the all-incoming convention define

$$\begin{aligned} -s &= (k_1 + k_2)^2, \\ -t &= (k_1 + k_3)^2, \\ -u &= (k_1 + k_4)^2. \end{aligned} \quad (6.9)$$

In the center-of-mass frame the spatial parts of $k_1 + k_2$ cancel, hence

$$s = E_{\text{cm}}^2. \quad (6.10)$$

For an outgoing physical momentum $q_3 = -k_3$,

$$k_1 + k_3 = k_1 - q_3 \quad (6.11)$$

is the momentum transferred from leg 1 to leg 3. For equal external masses, with incoming spatial momentum magnitude p ,

$$-t = 2p^2(1 - \cos \theta), \quad -u = 2p^2(1 + \cos \theta). \quad (6.12)$$

Thus t and u carry the angular information measured by detectors.

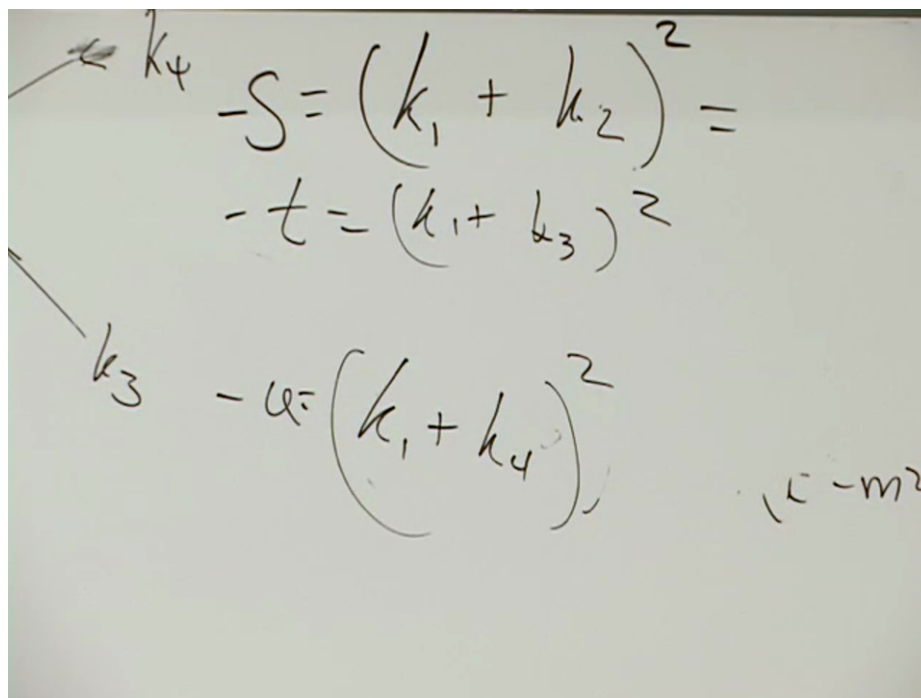


Figure 6.4: The symmetric definitions of s, t, u in the all-incoming convention.

Lecture frame: 00:26:20.

Three symbols do not mean three independent data. Squaring (6.7) and using $k_i^2 = -m^2$ gives

$$s + t + u = 4m^2 \quad (6.13)$$

for equal masses. Once s and t are known, u is fixed.

Question from the audience. Could the vacuum slow a freely moving particle, rather like friction? ^a

Response. Not in a translation-invariant vacuum. There is nowhere for the missing momentum to go, so momentum conservation forbids such friction. One can imagine a medium or a different vacuum that selects a preferred rest frame, but that is not the vacuum being used here.

^aLecture timestamp: 00:27:17.

Question from the audience. Why must the two outgoing particles have equal energies? ^a

Response. In the center-of-mass frame their spatial momenta must be equal and opposite. For the equal-mass process under discussion, equal momentum magnitudes give equal energies through $E^2 = p^2 + m^2$. If the outgoing masses differed, that last conclusion would have to be modified.

^aLecture timestamp: 00:27:54.

6.4 Direct and Crossed Channels

A simple point-particle diagram lets particles 1 and 2 form an intermediate particle of mass M , which then decays into 3 and 4. Up to numerator factors and the pole prescription, its characteristic dependence is

$$\mathcal{A}_s \sim \frac{g^2}{s - M^2}. \quad (6.14)$$

This is the direct, or s -channel, process. Turning the connectivity sideways gives exchange between the two external lines,

$$\mathcal{A}_t \sim \frac{g^2}{t - M^2}, \quad (6.15)$$

the crossed, or t -channel, process. A third connectivity gives

$$\mathcal{A}_u \sim \frac{g^2}{u - M^2}. \quad (6.16)$$

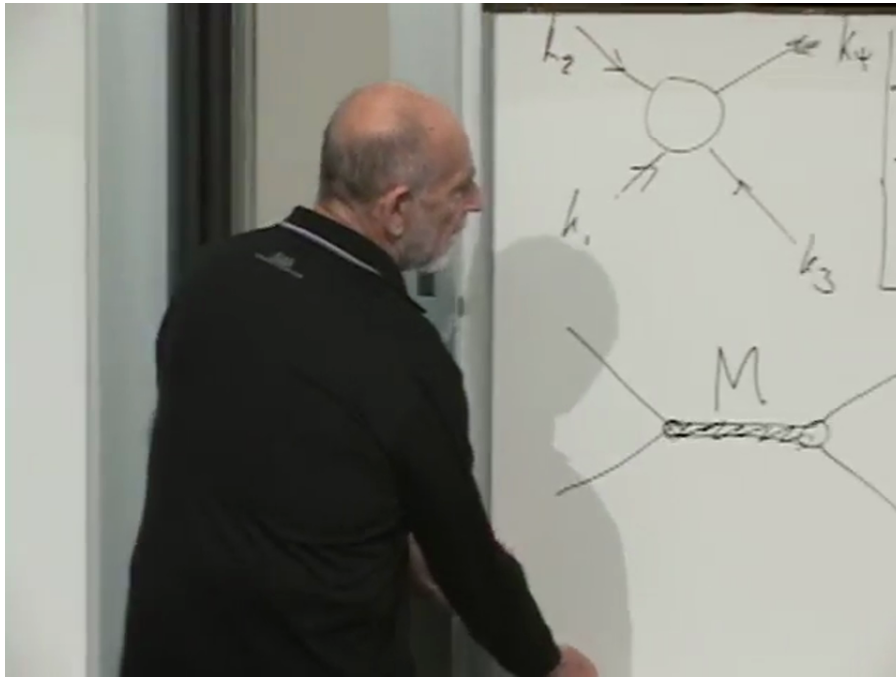


Figure 6.5: Direct production through an intermediate state and the corresponding pole $g^2/(s - M^2)$; the crossed connectivity is being assembled on the same board.

Lecture frame: 00:30:50.

The distinction is physical. At this crude scalar level, (6.14) depends on collision energy but not on angle. The compound state has forgotten the incident direction before it decays, so its decay is isotropic. The exchange amplitude depends on momentum transfer and therefore on angle; a light exchanged particle strongly favors small deflection. Spin would add further angular dependence, but the pole structure is already visible.

Question from the audience. Does the missing u -channel produce a third term? ^a

Response. Yes, whenever the particle identities and couplings permit that connectivity. One then includes a term with a pole at $u = M^2$. It is not a new independent kinematic variable, because $s + t + u = 4m^2$, but it is a distinct channel diagram.

^aLecture timestamp: 00:34:47.

6.5 One Function, Two Channel Descriptions

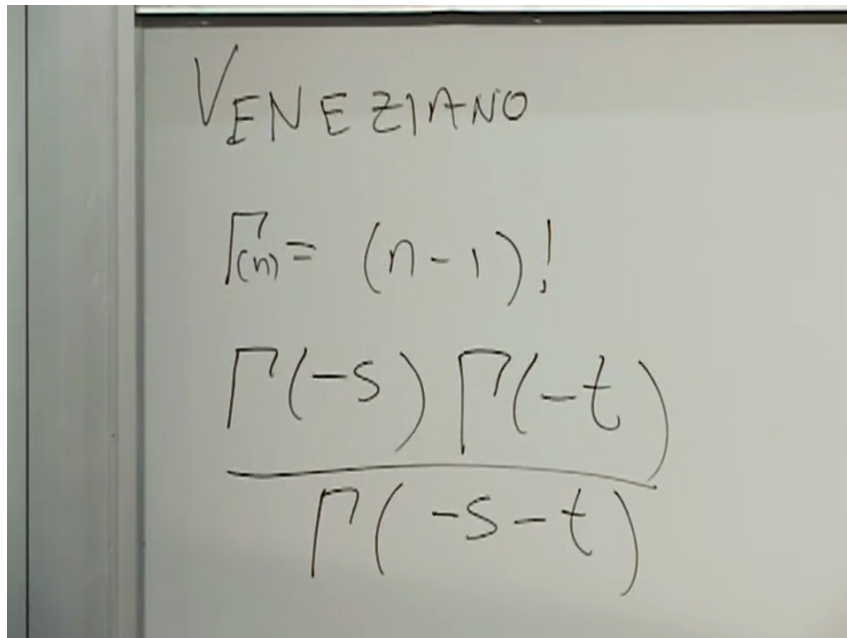
A single pole is a poor model of meson scattering. Two mesons can form many resonances, with a tower of masses and angular momenta. The first response in the 1960s was to add more s -channel poles, more t -channel poles, and the allowed u -channel poles. Then a much more economical function was found. In the lecture's units, where the linear Regge trajectory and its intercept have been absorbed into s and t , it is

$$\mathcal{A}_V(s, t) \propto \frac{\Gamma(-s)\Gamma(-t)}{\Gamma(-s-t)}. \quad (6.17)$$

The gamma function continues the factorial:

$$\Gamma(n) = (n-1)! \quad (6.18)$$

at positive integers.



The image shows a whiteboard with the following handwritten text:

$$VENEZIANO$$

$$\Gamma(n) = (n-1)!$$

$$\frac{\Gamma(-s)\Gamma(-t)}{\Gamma(-s-t)}$$

Figure 6.6: The Veneziano amplitude written as a symmetric ratio of gamma functions.

Lecture frame: 00:42:25.

The poles of $\Gamma(-s)$ reproduce an infinite tower of direct-channel resonances. Yet the same expression is symmetric under $s \leftrightarrow t$, so its $\Gamma(-t)$ poles give a crossed-channel expansion. These are not two separately added contributions. One analytic function admits either description. This is the old idea of duality between direct-channel resonances and crossed-channel exchanges.

Question from the audience. How can one amplitude contain both channel sums without adding them? ^a

Response. That is precisely the surprise. Expanding the function near its s -channel poles organizes it as direct resonances; expanding it near its t -channel poles organizes the same function as exchanges. A microscopic model is needed to explain why those two readings are the same, and the model is an interacting string.

^aLecture timestamp: 00:43:32.

6.6 A Strip Swept Out by an Open String

Parameterize an open string by

$$0 \leq \sigma \leq \pi. \quad (6.19)$$

The label σ is not a spacetime position. The embedding $X^\mu(\sigma, \tau)$ tells us where that labeled point lies in spacetime, while τ measures evolution. A point particle sweeps out a worldline; the moving string sweeps out a two-dimensional worldsheet. In the free theory each embedding coordinate obeys

$$\frac{\partial^2 X^\mu}{\partial \tau^2} - \frac{\partial^2 X^\mu}{\partial \sigma^2} = 0. \quad (6.20)$$

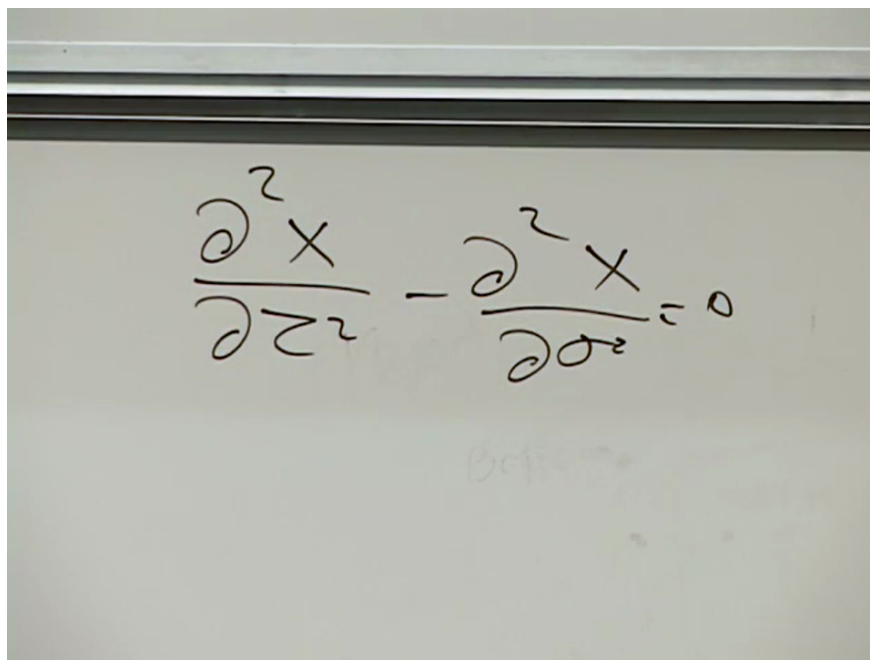


Figure 6.7: The wave equation governing each free string coordinate on the worldsheet strip.

Lecture frame: 00:49:10.

For the scattering construction it is useful not to take the continuum limit too soon. Replace each string by n mass points joined by springs. The worldsheet becomes a dense collection of nearby worldlines, and its normal modes are ordinary harmonic oscillators. Two incoming chains can touch at their endpoints. At that event a new spring joins them, they evolve for a time τ as one longer chain, and later a spring breaks so that two outgoing chains emerge.

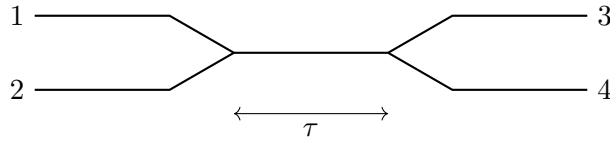
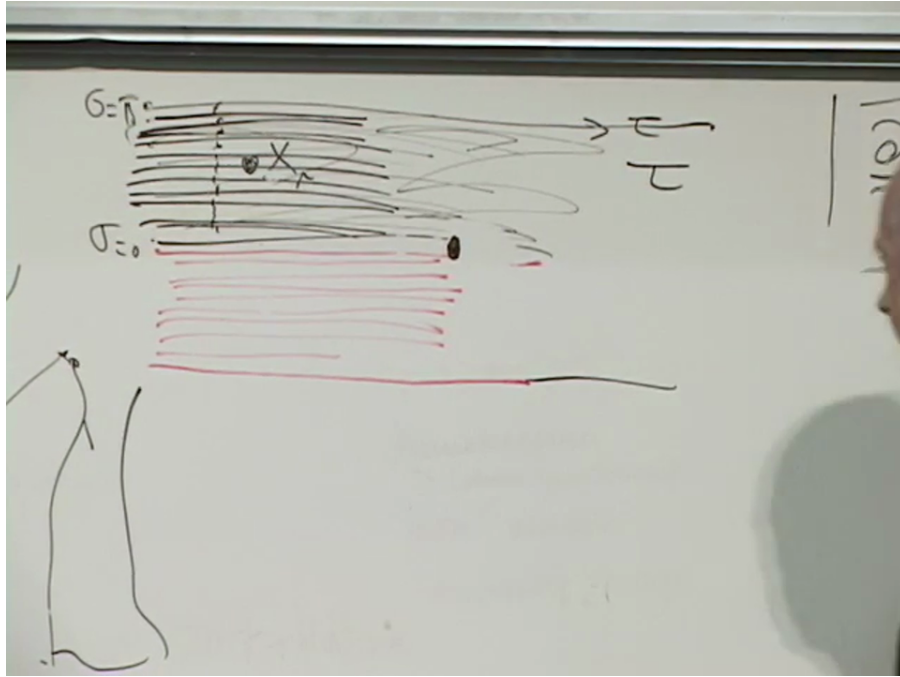


Figure 6.8: Two discretized open strings touch, form one intermediate string for a duration τ , and split. The source frame preserves the spring-and-worldline construction; the schematic isolates its connectivity.

Lecture frame: 00:51:25.

The two strings are drawn adjacent in parameter space only. Their spacetime positions need not be adjacent until the endpoint coordinates actually coincide. Reparameterization freedom lets the joining and splitting events be placed at convenient positions on the diagram; this geometric freedom will become conformal symmetry in the next lecture.

6.7 The Oscillator Calculation in Slow Motion

Let the first chain have coordinates $x_1, \dots, x_n$, and the second $x_{n+1}, \dots, x_{2n}$. Their center-of-mass coordinates are

$$X_{\text{cm}}^{(1)} = \frac{1}{n} \sum_{a=1}^n x_a, \quad X_{\text{cm}}^{(2)} = \frac{1}{n} \sum_{a=n+1}^{2n} x_a. \quad (6.21)$$

A chain with definite momentum k_1 has a center-of-mass plane wave multiplied by an internal oscillator ground state:

$$\Psi_1 = e^{ik_1 \cdot X_{\text{cm}}^{(1)}} \psi_0(x_1, \dots, x_n), \quad (6.22)$$

$$\Psi_2 = e^{ik_2 \cdot X_{\text{cm}}^{(2)}} \psi_0(x_{n+1}, \dots, x_{2n}). \quad (6.23)$$

The factors ψ_0 depend on relative coordinates, not on the common translation of an entire chain. Because the incoming strings are initially separate quantum systems,

$$\Psi_{\text{in}} = \Psi_1 \Psi_2. \quad (6.24)$$

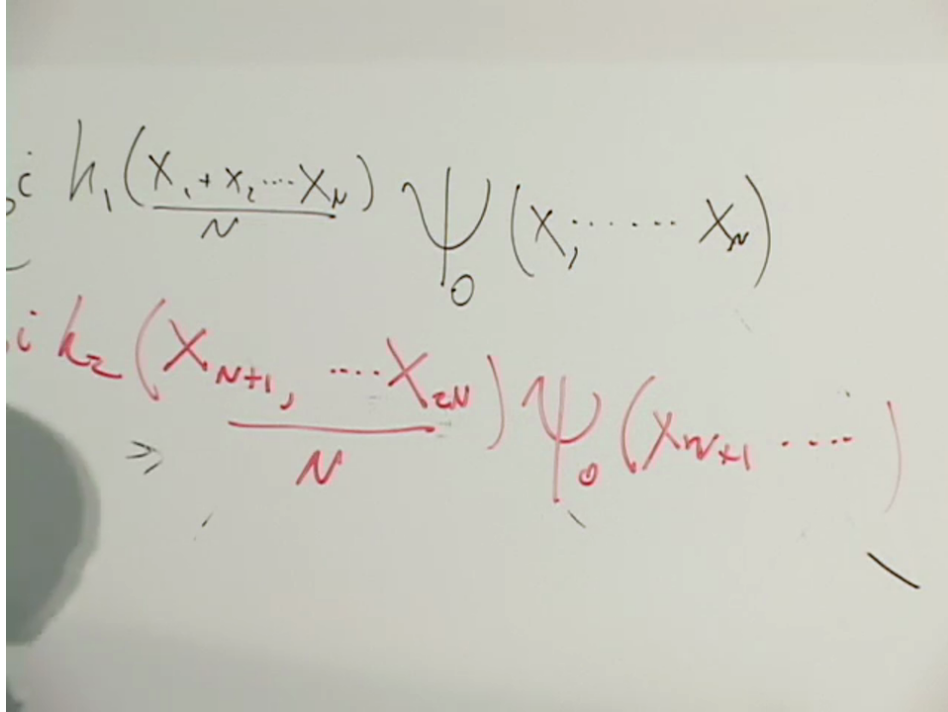


Figure 6.9: The two momentum plane waves and their internal oscillator ground-state factors, written in separate colors.

Lecture frame: 00:58:05.

Joining selects the part of this wavefunction for which the two endpoint coordinates agree:

$$x_n = x_{n+1}. \quad (6.25)$$

At that instant the other coordinates are unchanged. The selected state is then evolved with the Hamiltonian H_{joined} of the single longer chain:

$$\Psi(\tau) = e^{-iH_{\text{joined}}\tau} \Psi(0). \quad (6.26)$$

Finally one takes its overlap with a product of outgoing string states carrying momenta q_3, q_4 .

$$X_N \dots X_2) = e^{i h_1(\frac{X_1 + X_2 + \dots + X_N}{N})} \psi_0(X_1, \dots, X_N)$$

$$e^{i h_2(\frac{X_{N+1} + \dots + X_{2N}}{N})} \psi_0(X_{N+1}, \dots, X_{2N})$$

Figure 6.10: The two-chain wavefunction immediately before the endpoint coincidence is imposed.

Lecture frame: 00:58:55.

The calculation is lengthy but not mysterious. Every internal coordinate can be decomposed into oscillator modes, so the joining overlap, time evolution, and final projection reduce to Gaussian oscillator algebra. What remains unfixed is the lifetime τ of the joined string.

Question from the audience. Are the black and red wavefunctions added or multiplied? ^a

Response. Multiplied. The tensor-product state of two independent systems is represented in the coordinate basis by the product of their wavefunctions. Addition would describe a superposition of alternatives, not the simultaneous presence of both incoming strings.

^aLecture timestamp: 01:02:26.

Question from the audience. Why integrate over the intermediate time τ ? ^a

Response. The joined string is not assigned a fixed lifetime. Feynman's rule is to sum amplitudes over all allowed histories. In this description the remaining continuous history parameter is the time between joining and splitting, so the sum becomes an integral over $0 \leq \tau < \infty$.

^aLecture timestamp: 01:04:04.

6.8 From Propagation Time to the Beta Function

After evaluating the oscillator overlaps, the result can be organized as

$$A(s, t) \propto \int_0^\infty d\tau e^{\tau(s+1)} (1 - e^{-\tau})^{-t-1} e^{-\tau}. \quad (6.27)$$

The final $e^{-\tau}$ is kept visible because it is the Jacobian needed in the next step. Equivalently, the first and last exponentials combine to $e^{s\tau}$.

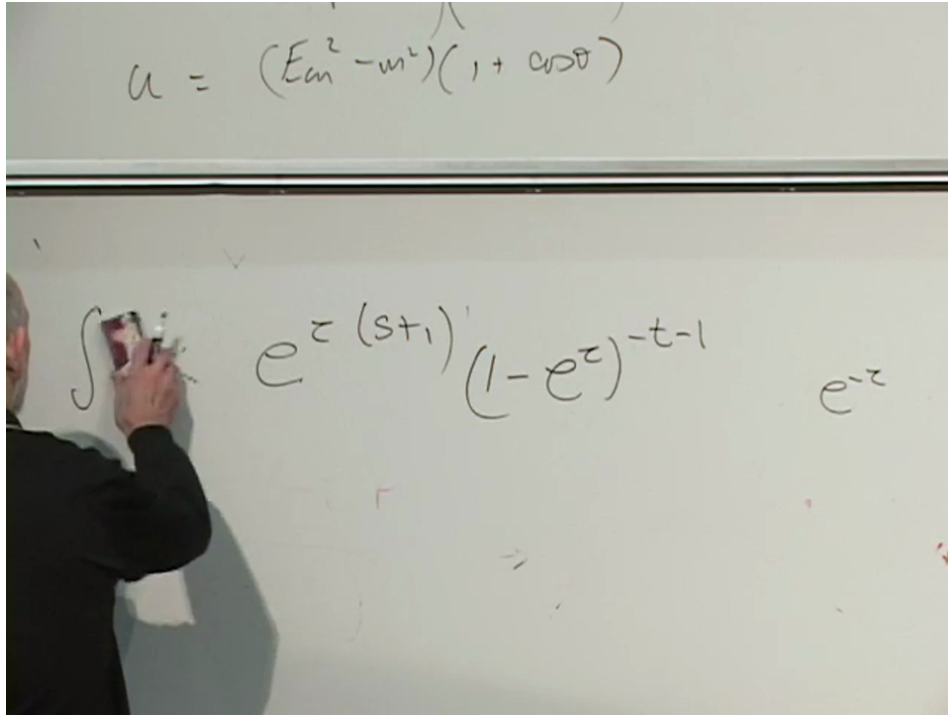


Figure 6.11: The propagation-time integrand after the oscillator calculation; the separate $e^{-\tau}$ factor supplies the change-of-variables Jacobian.

Lecture frame: 01:06:45.

Now set

$$z = e^{-\tau}, \quad dz = -e^{-\tau} d\tau. \quad (6.28)$$

As τ runs from 0 to ∞ , z runs from 1 to 0. The minus sign reverses the limits, giving

$$\mathcal{A}(s, t) \propto \int_0^1 z^{-s-1} (1-z)^{-t-1} dz. \quad (6.29)$$

The image shows a handwritten derivation on a chalkboard. It starts with the integral $\int_0^\infty e^{-\tau} (s-1)' (1-e^{-\tau})^{-t-1}$. Below this, the substitution $e^{-\tau} = z$ is written. The final integral is $\int z^{s+1} (1-z)^{-t-1} dz$.

Figure 6.12: The substitution $z = e^{-\tau}$ exposes the Euler beta integral.

Lecture frame: 01:08:49.

This integral is defined directly where its endpoints converge and elsewhere by analytic continuation. It is the Euler beta function:

$$B(-s, -t) = \frac{\Gamma(-s)\Gamma(-t)}{\Gamma(-s-t)}. \quad (6.30)$$

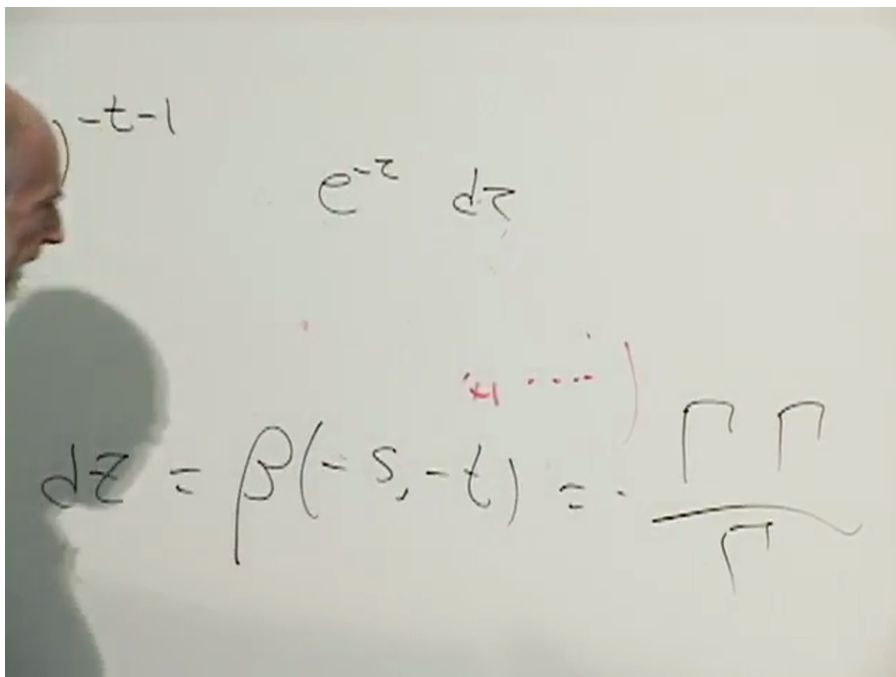


Figure 6.13: The board identifies the transformed integral with $B(-s, -t)$ and begins restoring its gamma-function form.

Lecture frame: 01:12:24.

The channel symmetry is now elementary. Under $z \mapsto 1 - z$,

$$\int_0^1 z^{-s-1} (1-z)^{-t-1} dz = \int_0^1 z^{-t-1} (1-z)^{-s-1} dz. \quad (6.31)$$

Thus $\mathcal{A}(s, t) = \mathcal{A}(t, s)$. A process constructed as two strings joining and later splitting automatically contains the crossed description as well.²

Question from the audience. The gamma-function ratio resembles factorial expressions in combinatorics. Is there a direct physical counting interpretation?^a

Response. At integer arguments the beta function is related to a ratio such as $n!m!/(n+m)!$, but no combinatorial mechanism is being asserted here. The same special function happens to arise from the worldsheet integral; this resemblance was not the route by which the amplitude was found.

^aLecture timestamp: 01:13:22.

The still-hidden reason why one worldsheet can be read in both channels is conformal symmetry. It permits the sheet to be stretched like taffy without changing the amplitude, turning a long direct-channel propagation region into a crossed-channel exchange picture. That geometric explanation is deferred to the next lecture.

6.9 What the Scattering Test Established

The discrete joining picture prompted a final cluster of questions.

²Editorial clarification: The lecture suppresses the Regge slope and intercept in the arguments of the beta function. In standard notation one replaces s, t by the corresponding linear trajectories $\alpha(s), \alpha(t)$. The duality and the change-of-variables argument are unchanged.

Question from the audience. If two chains of n constituents join at their endpoints, is the result $2n - 1$, and could that change fermion parity? ^a

Response. In the hadronic endpoint picture the touching endpoint pair may be regarded as a particle-antiparticle pair that annihilates, so the count is reduced by two, not one. In the continuum limit the finite difference between $2n$ and $2n - 2$ is immaterial, but preserving fermion parity is not. The lattice constituents are only a regulator for the continuous string, not literal fundamental beads.

^aLecture timestamp: 01:16:16.

Question from the audience. Will the later discussion continue into the full mathematics of superstrings? ^a

Response. Only selected consequences will be used. The immediate route is through worldsheet symmetry and compactification, followed by a different starting point developed in the mid-1990s: M-theory. Its purpose is to reach the same physics from a viewpoint in which several otherwise complicated features become transparent.

^aLecture timestamp: 01:18:05.

Question from the audience. Was this joining rule only a clue from meson collisions? ^a

Response. Meson scattering supplied the historical clue, but the rule is general for open strings. The same topology describes collisions of open-string photons and other open-string states. Two closed strings join and split by an analogous construction, with a closed string in the intermediate region.

^aLecture timestamp: 01:20:02.

Question from the audience. Does the calculation apply only when the string states are fermions? ^a

Response. No. The original Veneziano calculation belongs to the bosonic open string. Superstrings require additional fermionic degrees of freedom and modify details of the integrand, but the joining, propagation, projection, and worldsheet-duality logic remains closely analogous.

^aLecture timestamp: 01:20:56.

Finding a spin-one or spin-two state in the oscillator spectrum was suggestive, but not decisive. Their scattering amplitudes supplied the stronger test. Photon amplitudes obey the identities associated with charge conservation and gauge invariance; graviton amplitudes obey the corresponding constraints of a massless spin-two field. The string states passed these tests in interactions with massive and massless states and with one another. That is why the photon and graviton labels became physical identifications rather than guesses based only on polarization counting. ³

Question from the audience. Where does electric charge itself come from in the string description? ^a

Response. That question is left open at the end of the lecture. Its answer requires the internal and endpoint structure of the string theory, which has not yet been introduced, so it belongs to the discussion that follows rather than to this scattering derivation.

^aLecture timestamp: 01:24:03.

³Editorial clarification: In modern language these tests are encoded in decoupling of unphysical polarizations, gauge Ward identities, and the universal low-energy coupling of the massless spin-two state.

Chapter 7

Worldsheets, Wick Rotation, and Channel Duality

Overview. A particle traces a worldline; a string sweeps out a worldsheet. Quantizing either object requires a sum over histories, but the Lorentzian sum is oscillatory and difficult to define. Turning worldsheet time into an imaginary coordinate gives a Euclidean theory governed by the Laplace equation. Its conformal symmetry then turns the slit surface of a four-string interaction into a disc with four boundary insertions, where the symmetry between direct and crossed channels becomes visible.

7.1 From a Worldline to a Sum over Histories

Begin with a free particle. In nonrelativistic mechanics its position is a function $\mathbf{x}(t)$, and, with the mass set to one, its action is

$$S_{\text{NR}}[\mathbf{x}] = \frac{1}{2} \int_{t_1}^{t_2} \dot{\mathbf{x}}^2 dt. \quad (7.1)$$

For a relativistic description it is more natural to place time and space in one vector, $X^\mu(\tau)$, parameterized along the trajectory by τ . The corresponding quadratic expression is

$$S_{\text{rel}}[X] = \frac{1}{2} \int_{\tau_1}^{\tau_2} \frac{dX^\mu}{d\tau} \frac{dX_\mu}{d\tau} d\tau. \quad (7.2)$$

The upper and lower indices encode the Lorentzian sign difference between time and space. ¹

¹Editorial clarification: The quadratic relativistic action is being used in a fixed worldline parameterization. The manifestly reparameterization-invariant alternative is proportional to proper length, which is mentioned later in the lecture.

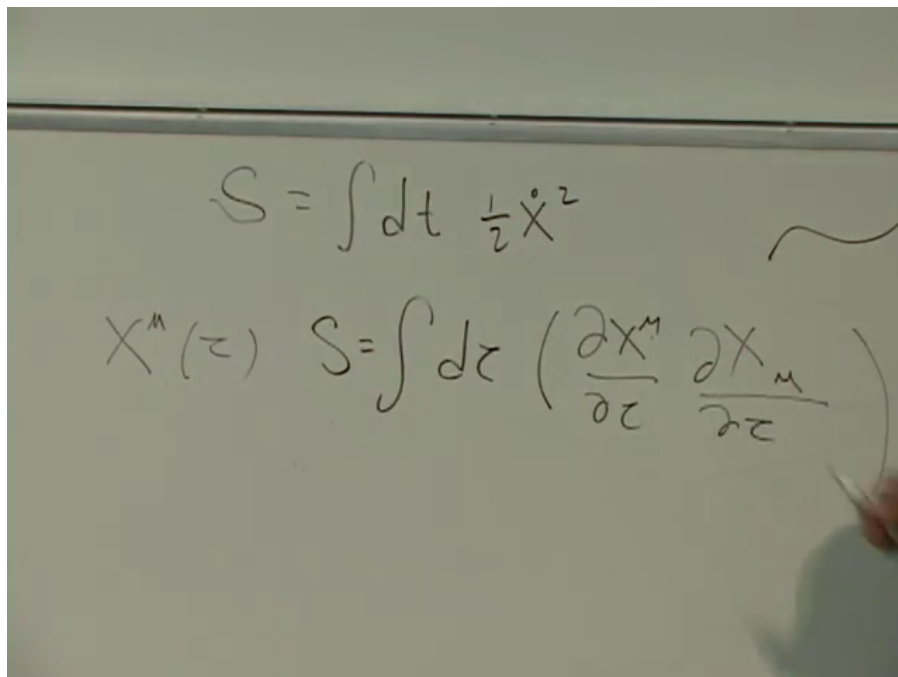


Figure 7.1: The nonrelativistic action and its relativistic quadratic counterpart written together on the board.

Lecture frame: 00:03:35.

Classically, one fixes the endpoints and varies the path. A free particle then follows a straight worldline with constant velocity. Quantum mechanics asks a different question. It does not select one trajectory between the endpoints; it asks for the amplitude to begin at one point and be detected at the other. Every intervening path contributes a phase,

$$K(2, 1) = \int_{X(1)}^{X(2)} \mathcal{D}X e^{iS[X]}. \quad (7.3)$$

The functional integral is an integral over an infinite-dimensional space of paths, not the ordinary integral already appearing inside the action.

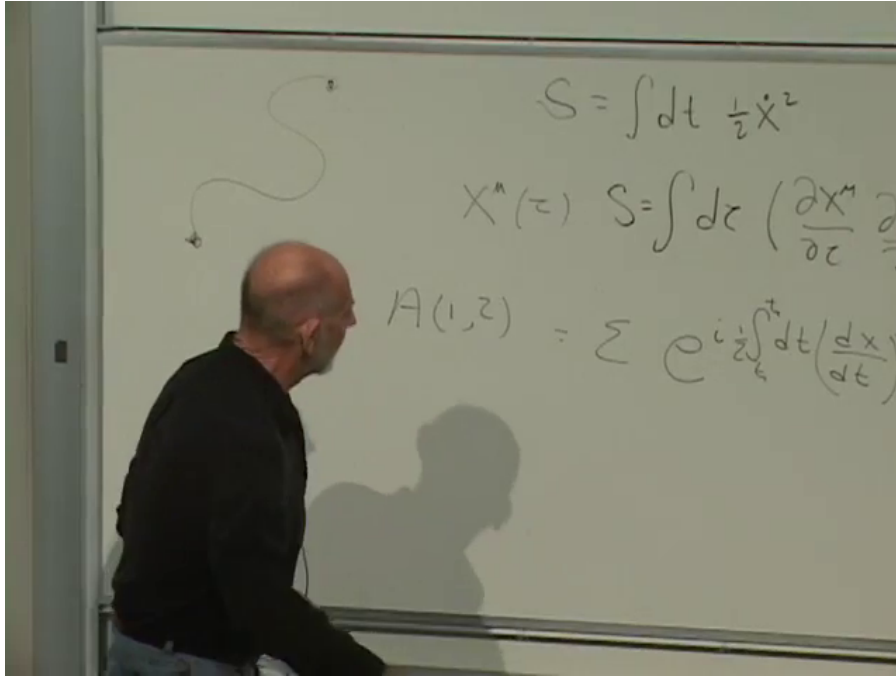


Figure 7.2: A representative path and the phase assigned to it before the sum over all trajectories is taken.

Lecture frame: 00:06:21.

In relativistic language this point-to-point amplitude is a propagator. For an interaction, several worldline segments meet. Their actions are added, the meeting points are integrated over, and diagrams with the required external particles are summed. Thus the familiar Feynman-diagram expansion is already implicit in the instruction to sum over histories.

The symbol S will denote an action here. It is unrelated both to entropy and to the Mandelstam variable s used near the end of the lecture. The shared notation is traditional rather than illuminating.

7.2 Turning Time Sideways

There is an immediate problem with (7.3). Every factor e^{iS} has magnitude one. A wildly oscillating trajectory and the classical straight line therefore enter with equal absolute weight; only their phases differ. Destructive interference recovers classical behavior, but it does not make the formal integral an ordinary convergent object.

Introduce a new parameter s through

$$\tau = \alpha s, \quad d\tau = \alpha ds, \quad \frac{dX}{d\tau} = \frac{1}{\alpha} \frac{dX}{ds}. \quad (7.4)$$

Then the quadratic action rescales as

$$\frac{1}{2} \int d\tau \left(\frac{dX}{d\tau} \right)^2 = \frac{1}{2\alpha} \int ds \left(\frac{dX}{ds} \right)^2. \quad (7.5)$$

Choosing the contour corresponding to $\alpha = -i$ changes the oscillatory weight into

$$e^{iS} \rightarrow \exp \left[-\frac{1}{2} \int ds \left(\frac{dX}{ds} \right)^2 \right] = e^{-S_E}. \quad (7.6)$$

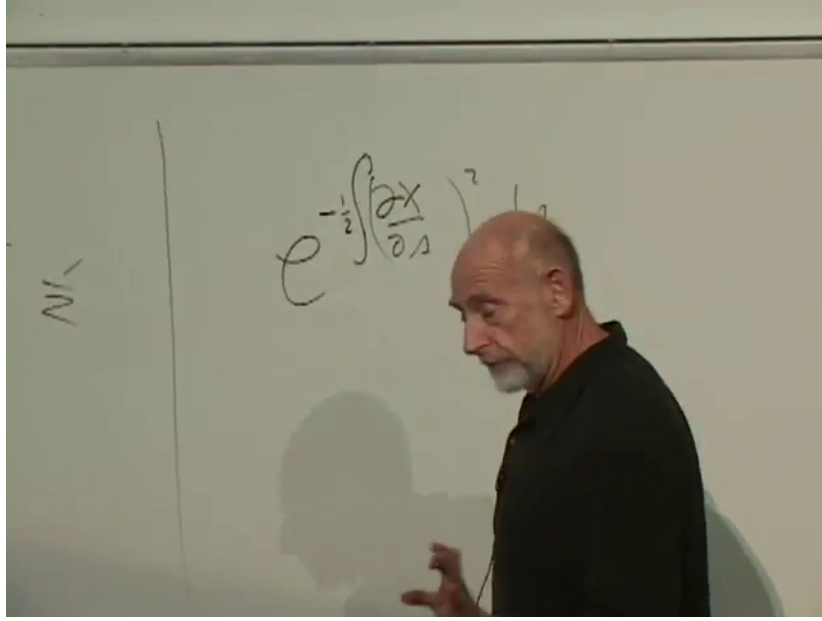


Figure 7.3: The Euclidean particle weight after the imaginary-time substitution; the exponent is real and nonpositive.

Lecture frame: 00:16:20.

A path that runs far away and returns, or wiggles rapidly without traveling far, has a large positive S_E and is exponentially suppressed. The result is a Wiener-type integral, for which a serious convergence theory exists. The endpoints inherited from real τ lie on an imaginary s -contour. In practice one evaluates the Euclidean object in its convergent domain and analytically continues the answer back to the physical Lorentzian variables. This operation is the Wick rotation: time has been turned sideways for the calculation, not discarded.

7.3 From Worldlines to Worldsheets

A point particle has a worldline. An open string sweeps out a ribbon, and a closed string sweeps out a tube. Coordinates (σ, τ) label points on this two-dimensional worldsheet, while

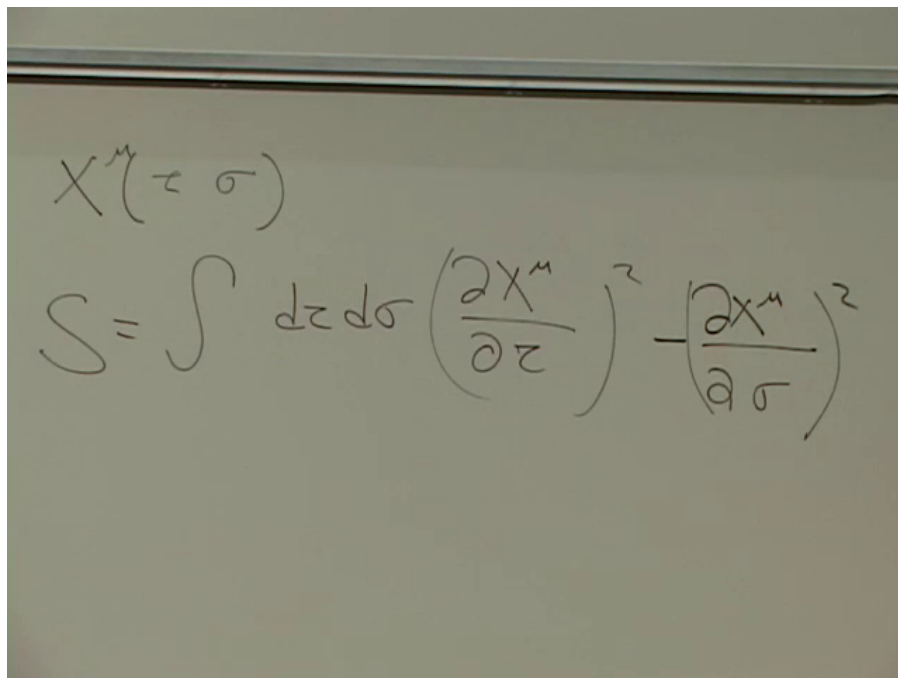
$$X^\mu = X^\mu(\sigma, \tau) \quad (7.7)$$

gives the spacetime position of the labeled point. Knowing all the functions X^μ specifies the complete history of the string.

The quadratic Lorentzian worldsheet action has the same kinetic-minus-potential structure as an ordinary vibrating string:

$$S_{\text{ws}}[X] = \frac{T_s}{2} \int d\tau d\sigma [\partial_\tau X^\mu \partial_\tau X_\mu - \partial_\sigma X^\mu \partial_\sigma X_\mu]. \quad (7.8)$$

The overall tension T_s fixes the scale; the lecture suppresses it while concentrating on the derivative structure.



A photograph of a whiteboard with handwritten mathematical equations. The top equation is $X^\mu(\tau, \sigma)$. Below it is the action $S = \int d\tau d\sigma \left(\frac{\partial X^\mu}{\partial \tau} \right)^2 - \left(\frac{\partial X^\mu}{\partial \sigma} \right)^2$. The handwriting is in dark ink on a light-colored surface.

Figure 7.4: The Lorentzian worldsheet action for $X^\mu(\sigma, \tau)$, with its kinetic-minus-potential sign.

Lecture frame: 00:24:26.

The string amplitude is obtained by exponentiating this action and summing over all surfaces connecting prescribed initial and final string configurations:

$$\mathcal{A} = \int \mathcal{D}X e^{iS_{\text{ws}}[X]}. \quad (7.9)$$

The same rule handles two strings becoming two, two becoming three, or any other allowed external configuration. Different worldsheet topologies also belong to the sum. Adding holes is the string counterpart of adding loop orders in a point-particle diagram expansion.

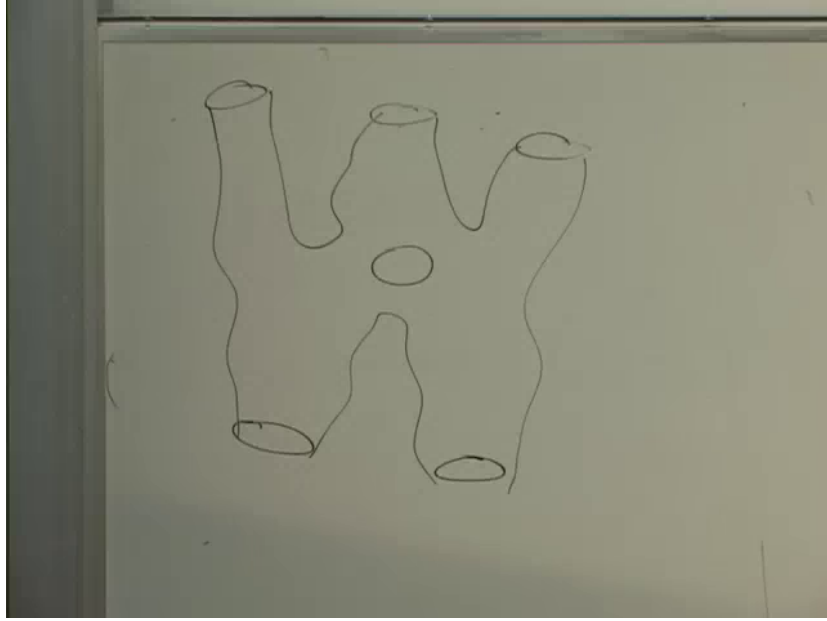


Figure 7.5: A surface connecting several external strings, with a hole representing a different worldsheet topology.

Lecture frame: 00:27:14.

Question from the audience. For a point electron, does the σ -derivative simply vanish? ^a

Response. A point particle has no σ coordinate in the first place. Nevertheless the pictures are connected: when external states are light compared with the string scale and process energies are low, string amplitudes reproduce ordinary point-particle Feynman diagrams. That limit is more than setting one derivative to zero.

^aLecture timestamp: 00:27:56.

The surface integral has the same convergence problem as the path integral. After Wick-rotating worldsheet time, its stable Euclidean form is

$$S_{E,ws}[X] = \frac{T_s}{2} \int d\tau d\sigma [\partial_\tau X^\mu \partial_\tau X_\mu + \partial_\sigma X^\mu \partial_\sigma X_\mu],$$

$$\mathcal{A}_E = \int \mathcal{D}X e^{-S_{E,ws}}.$$
(7.10)

Both derivatives now enter with the same positive sign. A violently vibrating or greatly stretched surface has large Euclidean action and is suppressed.

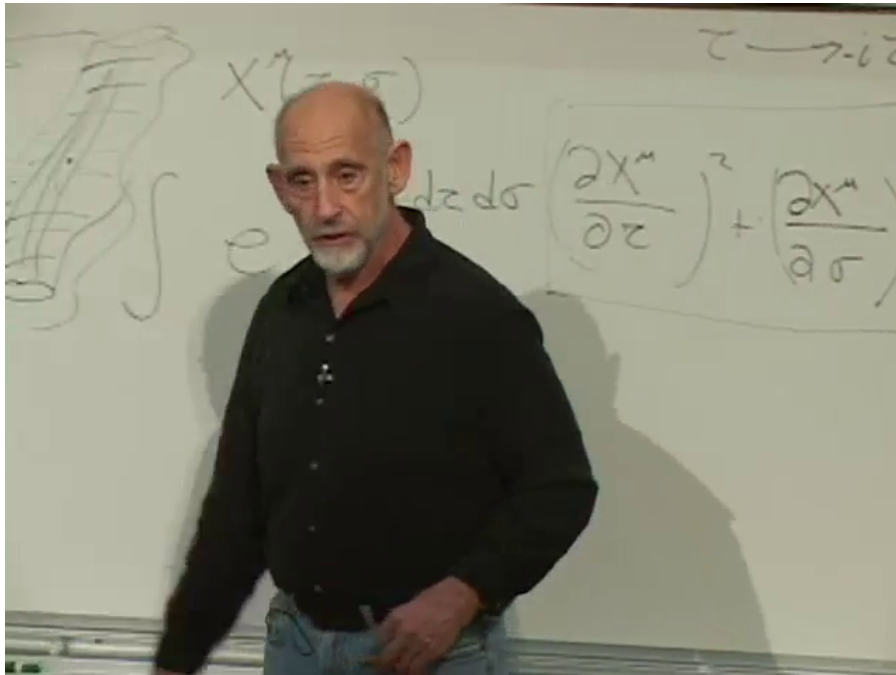


Figure 7.6: The positive Euclidean worldsheet density after τ is rotated to imaginary time.

Lecture frame: 00:34:29.

There is also a geometrically direct formulation. A relativistic particle may be assigned an action proportional to the proper length of its path; a string may be assigned an action proportional to the spacetime area of its worldsheet. The latter is the Nambu–Goto action. The quadratic worldsheet form used here is conventionally called the Polyakov form and is much easier for the present calculation.²

7.4 The Euclidean Equation of Motion

The coordinates σ and τ are only labels on the sheet. The next question is therefore unavoidable: which relabelings leave the theory unchanged? The answer is easiest to see from the equation of motion. Varying (7.10) gives, component by component,

$$\partial_\tau^2 X^\mu + \partial_\sigma^2 X^\mu = 0. \quad (7.11)$$

Before Wick rotation the relative sign is negative and this is a wave equation. In Euclidean signature both signs are positive, so it is the two-dimensional Laplace equation.

Question from the audience. Is this the equation of a vibrating drum?^a

Response. It is closely related, but a vibrating drum also carries a real time coordinate and therefore obeys a wave equation. The equation here is Euclidean Laplace’s equation. Electrostatic potential in a charge-free region supplies a more direct familiar example.

^aLecture timestamp: 00:41:50.

²Editorial clarification: The lecture includes a humorous discussion of historical naming and also mentions the related Liouville action and Dirichlet boundary conditions. No priority claim is needed for the derivation: only the equivalence of the area and auxiliary-metric formulations is used.

Question from the audience. The action contains first derivatives. Why does its equation contain second derivatives? ^a

Response. Varying a term quadratic in a first derivative and integrating by parts moves one derivative onto that first derivative. It is the field-theory version of $\dot{x}^2$ in an action producing $\ddot{x}$ in the equation of motion.

^aLecture timestamp: 00:42:44.

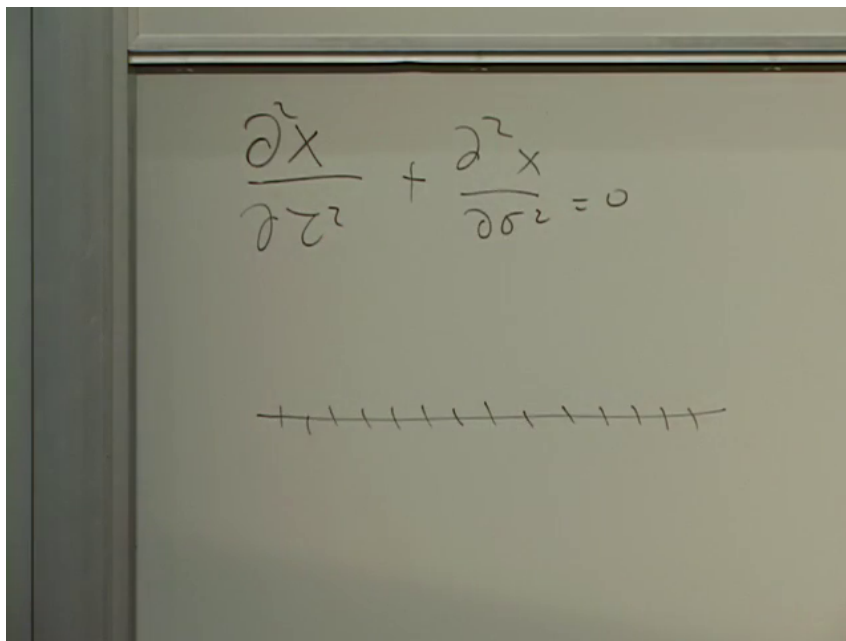


Figure 7.7: The plus-sign Laplace equation and the first discrete line used to interpret its second derivatives.

Lecture frame: 00:43:30.

To see the geometry, first discretize one coordinate with spacing ϵ . A first derivative is a neighboring difference, while a second derivative is a difference of neighboring differences:

$$\left. \frac{d^2 X}{ds^2} \right|_2 \longrightarrow \frac{X(3) - 2X(2) + X(1)}{\epsilon^2}. \quad (7.12)$$

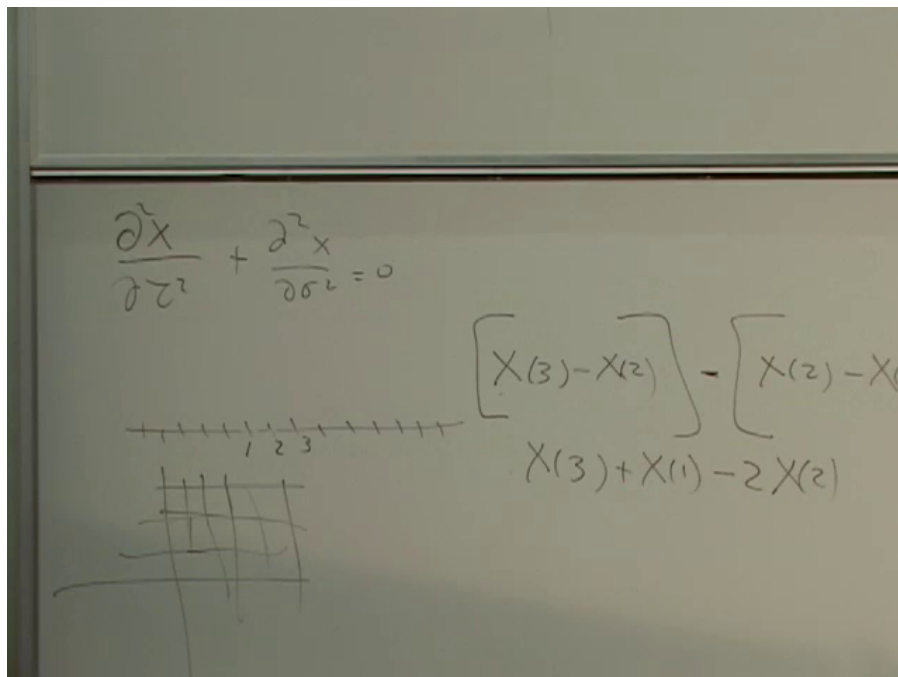


Figure 7.8: The numbered line, small lattice patch, and finite-difference expansion of a second derivative.

Lecture frame: 00:45:51.

Now discretize the (σ, τ) -plane. Label four nearest neighbors of a central point 5 by 1, 2, 3, 4. Equation (7.11) becomes

$$X(1) + X(2) + X(3) + X(4) - 4X(5) = 0, \quad (7.13)$$

or

$$X(5) = \frac{1}{4}[X(1) + X(2) + X(3) + X(4)]. \quad (7.14)$$

The image shows a handwritten derivation of the 2D finite-difference equation. At the top, there are two expressions in brackets: $[X(3) - X(2)]$ and $[X(2) - X(1)]$. Below these, the equation $X(3) + X(1) - 2X(2)$ is written. Further down, the equation $X(1) + X(3) - 2X(5) + X(2) + X(4) - 2X(5)$ is shown. Finally, the completed equation is written as $X(1) + X(2) + X(3) + X(4) = 4X(5)$. To the left of the equations, there are some faint markings, including a horizontal line with tick marks and the numbers 2 and 3.

Figure 7.9: The completed two-dimensional finite-difference equation: the central value equals the average of its four neighbors.

Lecture frame: 00:48:26.

This mean-value property holds for an infinitesimal square in any orientation. It is the local geometric content of the Laplace equation and the bridge to conformal symmetry.

7.5 Conformal Coordinates

Consider a change from (σ, τ) to (σ', τ') . A general deformation sends a small square to an arbitrary quadrilateral. The special transformations relevant here send every infinitesimal square to another infinitesimal square, possibly rotated and rescaled. Equivalently, they preserve the angles at which curves meet. These are conformal transformations.

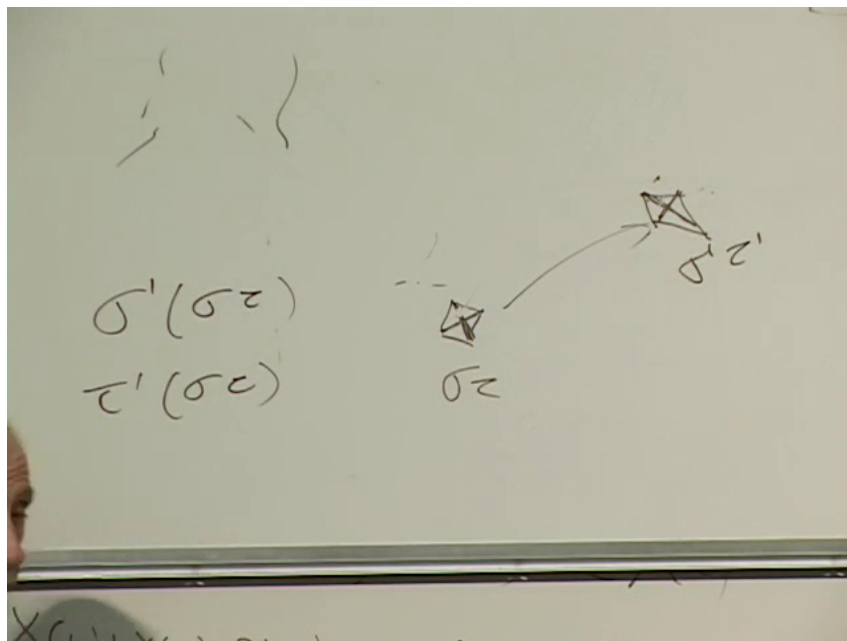


Figure 7.10: A local square mapped to another small figure while its angles are tested for conformal preservation.

Lecture frame: 00:53:25.

Question from the audience. If angles are preserved, is the transformation basically only a rotation? ^a

Response. Only locally, and even there a common scale change is also allowed. The local rotation and scale factor may vary from point to point, so the coordinate grid can bend, expand in one region, and contract in another while preserving infinitesimal shapes.

^aLecture timestamp: 00:53:31.

An angle-preserving map need not preserve size. If size were also fixed everywhere, the possibilities would collapse to the much smaller rigid class of translations and rotations. Conformal maps instead preserve local shape. This is why a map projection can keep the shape of a small island while greatly distorting its area.

Question from the audience. Is a conformal mapping equivalent to saying that the coordinate system is flat? ^a

Response. No. Flatness and conformality are different statements. A conformal map may use curved coordinate lines; what it preserves is the angle between intersecting curves. A Mercator-type map projection is a familiar example of local shape preservation with substantial scale distortion.

^aLecture timestamp: 00:55:35.

In two dimensions the required transformations are generated locally by analytic functions of a complex coordinate

$$z = \sigma + i\tau. \quad (7.15)$$

Finite boundaries can change appearance dramatically. A square region may be mapped to a smooth region, with the original corners sent to ordinary points on the new boundary rather than to its

visually prominent points. There is no contradiction: conformality is a local interior condition, and the map preserves infinitesimal angles wherever it is regular.

Question from the audience. Is this coordinate freedom the reason the earlier join-split diagram could be chosen symmetrically? ^a

Response. Yes. The same conformal freedom that preserves the worldsheet equations lets us replace the awkward interaction strip by a more symmetric domain. The remaining freedom does not remove the physical join-split parameter; it repackages that parameter as the relative position of boundary insertions.

^aLecture timestamp: 00:56:32.

7.6 The Slit Sheet and the Disc

Return to four-string scattering. Two open strings enter, their endpoints join, the combined string propagates, and it later splits. Drawn as a flat worldsheet, the process is an infinite strip slit from each end. The separation τ_0 between the tips of the slits is the lifetime of the intermediate string, and the amplitude integrates over it.

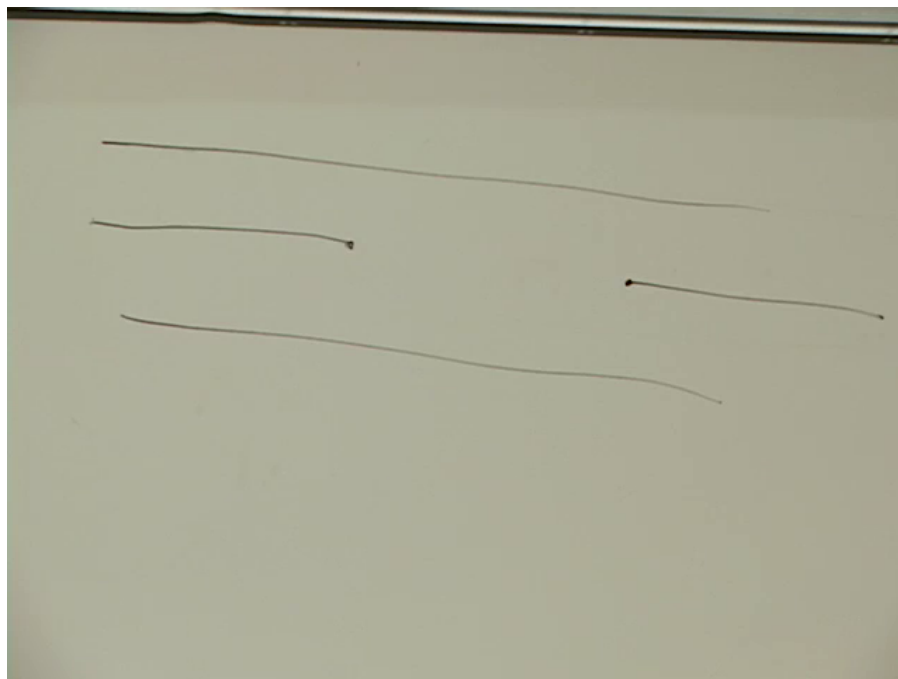


Figure 7.11: The doubly slit strip representing two open strings joining, propagating together, and splitting again.

Lecture frame: 01:00:20.

After Euclideanization, the embedding fields on this region satisfy the Laplace equation. Conformal invariance therefore permits the slit strip to be mapped to a disc. One way to see the topology is to foliate the disc by concentric circles and continuously deform those closed curves until they fill the slit region. The ends at infinity become exceptional boundary points under the map.

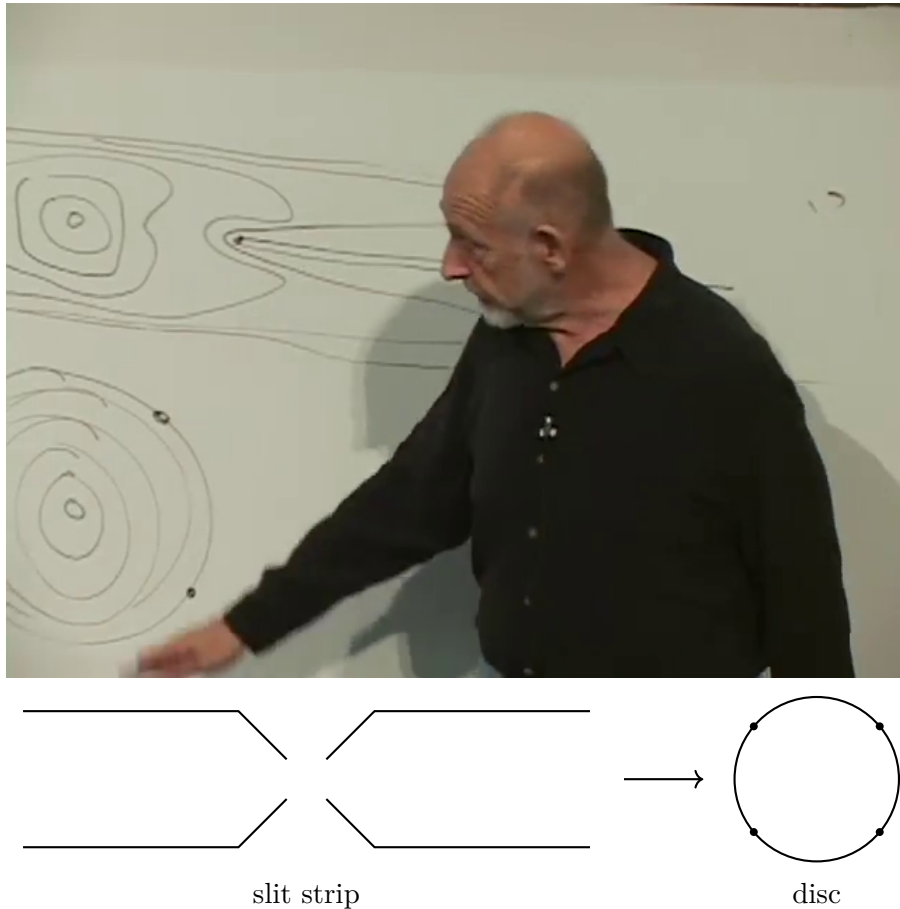


Figure 7.12: The board's level-curve comparison and a clean schematic of the conformal map from the slit strip to a disc. The four external ends become four special boundary points.

Lecture frame: 01:04:46.

The external states are not lost in this change of coordinates. Momentum, particle type, and oscillator level must be inserted at the four exceptional points. The formal objects carrying that information are vertex operators. Their detailed construction is deferred, but their role is already clear: they turn an otherwise homogeneous worldsheet path integral into a specified scattering experiment.

After using conformal transformations of the disc to move three boundary points to convenient reference positions, one real relative position remains. It may be represented by the cross-ratio of the four ordered boundary points. This is the disc version of the old lifetime τ_0 .

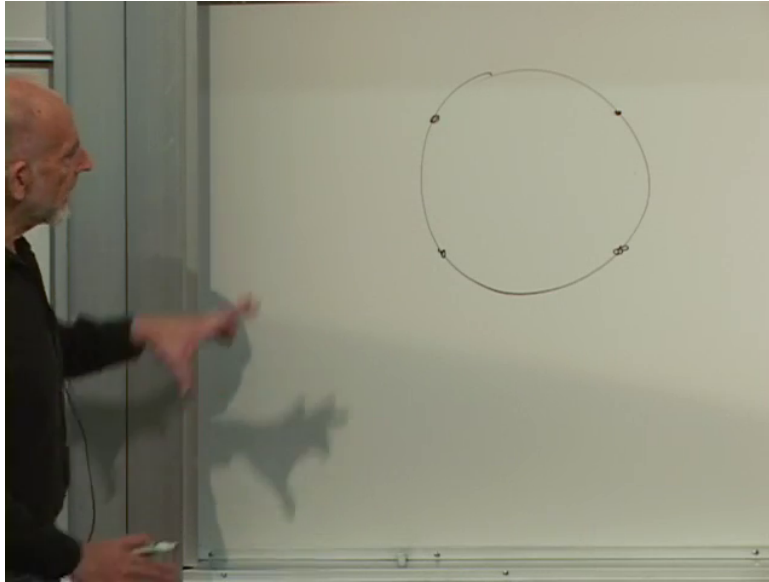


Figure 7.13: Four insertion points on the boundary of the disc. Conformal freedom removes all but one real modulus of their ordered configuration.

Lecture frame: 01:10:55.

The lifetime integral has therefore not disappeared. It has become an integral over that remaining modulus. With a conventional coordinate $0 < z < 1$, the four-point open-string integral takes the same form found in the preceding lecture,

$$\mathcal{A}(s, t) \propto \int_0^1 z^{-s-1} (1-z)^{-t-1} dz. \quad (7.16)$$

7.7 Two Limits of One Surface

At one end of moduli space, the two insertion points associated with the incoming pair approach one another. In the slit-strip coordinates this is a long interval between joining and splitting: a direct-channel intermediate string propagates for a long time. At the other end, a different neighboring pair approaches. The same surface is then naturally read as a crossed-channel exchange. These descriptions look unrelated on the strip but are neighboring boundary limits of one disc integral.

Returning to the mostly-plus convention of the previous lecture,

$$-s = (k_1 + k_2)^2, \quad -t = (k_1 + k_3)^2. \quad (7.17)$$

The direct limit organizes poles in s ; the crossed limit organizes poles in t .³

³Editorial clarification: Near the end of the lecture the board uses a shorthand sign for the Mandelstam invariants while also identifying s with positive center-of-mass energy squared. Equation (7.17) keeps the mostly-plus convention established in Lecture 6.

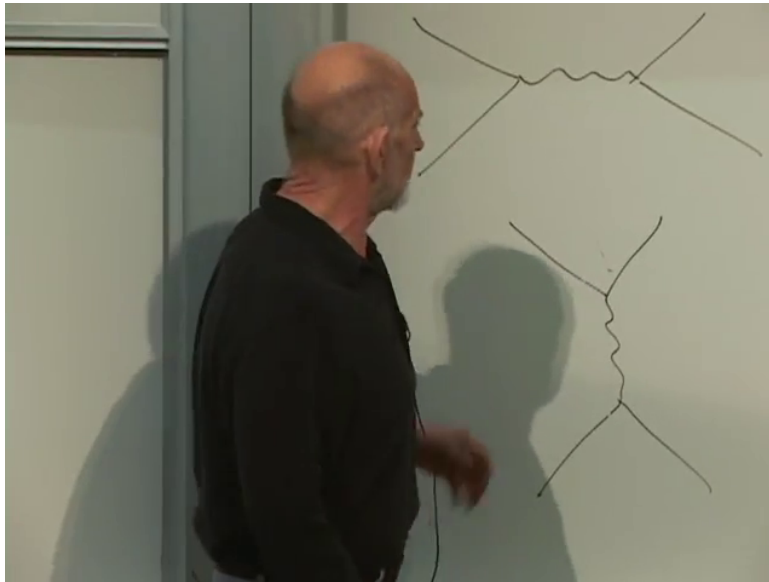


Figure 7.14: The two channel diagrams: long propagation in one orientation and exchange in the crossed orientation.

Lecture frame: 01:19:27.

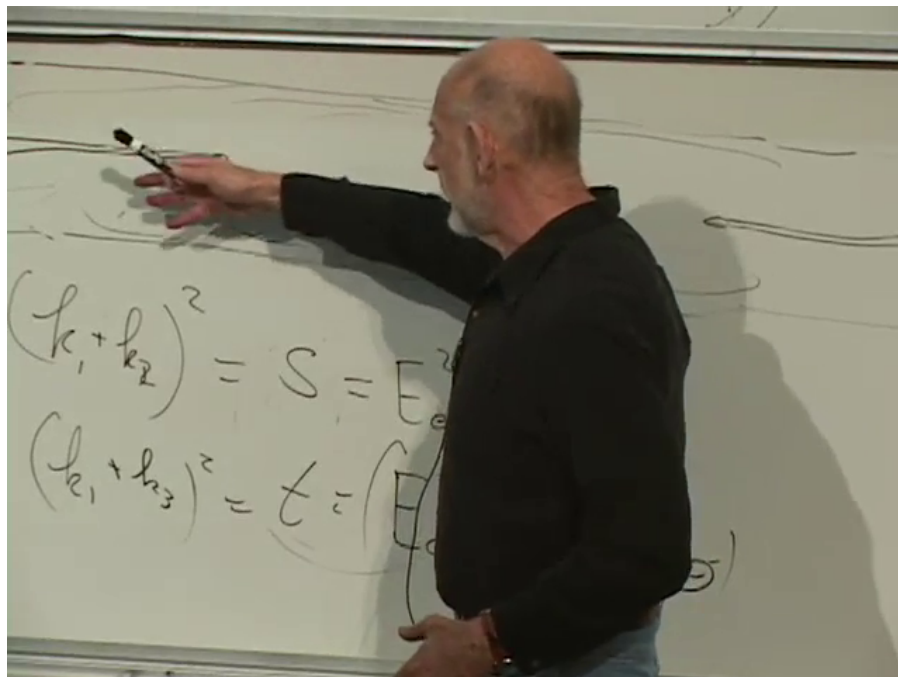


Figure 7.15: The board returns to the s - and t -channel momentum pairings while comparing the two limits of the disc modulus.

Lecture frame: 01:19:40.

One analytic object therefore supplies both expansions:

$$\mathcal{A}(s, t) = \mathcal{A}(t, s). \quad (7.18)$$

This is channel duality. In the original resonance language it appeared miraculous that a sum over direct-channel states could also encode crossed- channel exchange. On the Euclidean worldsheet, conformal symmetry explains why the two descriptions belong to the same surface. Historically this interpretation followed the 1969 discovery of the Veneziano amplitude very quickly.

Question from the audience. Is there a slit-strip picture corresponding to the crossed orientation as well? ^a

Response. Yes, but the direction called time must be chosen differently. The same Euclidean surface can be sliced so that another pair of external ends lies in the remote past and future. That alternative slicing is precisely why a single worldsheet admits more than one channel interpretation.

^aLecture timestamp: 01:20:49.

Question from the audience. Is the time for which the strings remain joined determined by the coupling constant? ^a

Response. No fixed lifetime is selected. The lifetime is an integration variable and every value contributes. The coupling controls the weight for joining and splitting—equivalently, the decay strength—and enters multiplicatively in the amplitude; it does not replace the modulus integral.

^aLecture timestamp: 01:21:39.

Chapter 8

Conformal Maps and the Open-String Scattering Prescription

Overview. Conformal mapping first appears as a practical symmetry of two-dimensional electrostatics. Complex differentiability turns that symmetry into a calculable one: the Cauchy–Riemann equations imply Laplace’s equation, and a direction-independent complex derivative preserves angles. With this machinery in hand, a string worldsheet may be drawn as a strip, half-plane, cylinder, or disc without changing the physics. Open-string scattering then becomes a Gaussian path integral on the disc with momentum insertions on its boundary, followed by an integral over their inequivalent positions.

8.1 Electrostatics in a Two-Dimensional World

We begin with a literal two-dimensional electrostatic problem. Charges and their field lines are confined to a plane; they are not merely ordinary three-dimensional charges whose motion happens to be planar. Gauss’s law then gives

$$F(r) \propto \frac{1}{r}, \quad V(r) \propto \log r, \quad (8.1)$$

because the flux crosses a circle whose circumference grows like r .

Write the field in terms of a scalar potential ϕ . In the sign convention used on the board,

$$\mathbf{E} = \nabla \phi, \quad \nabla \cdot \mathbf{E} = \rho, \quad (8.2)$$

and therefore

$$\nabla^2 \phi = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = \rho. \quad (8.3)$$

In a charge-free region this reduces to Laplace’s equation,

$$\nabla^2 \phi = 0. \quad (8.4)$$

1

¹Editorial clarification: Many electrodynamics texts define $\mathbf{E} = -\nabla \Phi$, in which case $\nabla^2 \Phi = -\rho$ after constants are absorbed. The lecture uses the opposite potential sign. The source-free equation and all conformal arguments are unchanged.

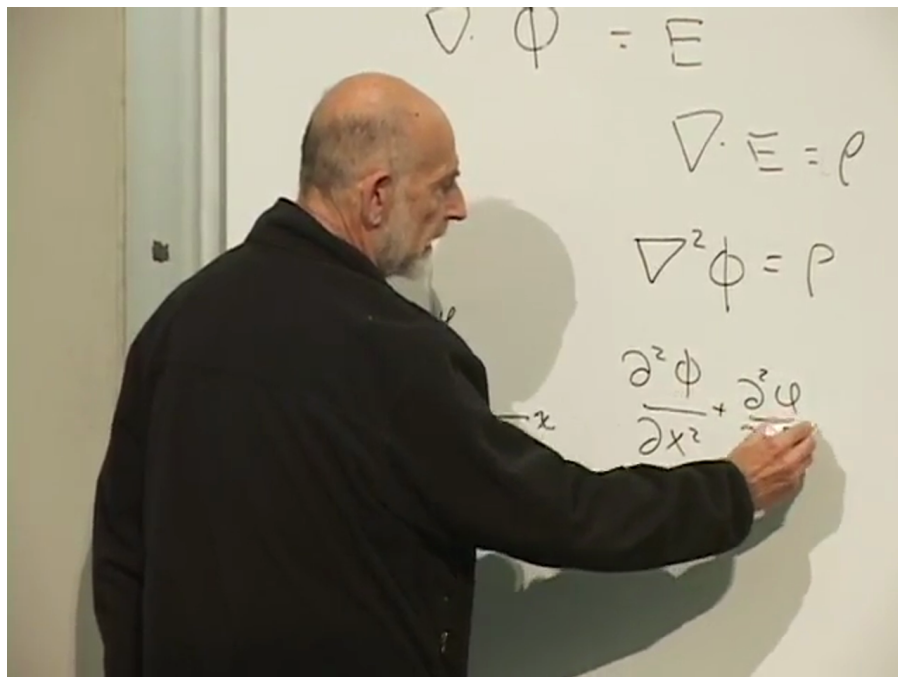


Figure 8.1: The electrostatic equations unfolded into the two-dimensional Poisson operator.

Lecture frame: 00:03:48.

The plus sign between the two spatial derivatives matters. It makes this an elliptic equation rather than a wave equation. Conformal transformations preserve its source-free form, so a solution on one domain may be carried to a differently shaped domain together with its boundary conditions. This is useful in electrostatics, two-dimensional fluid flow, and cartography before it is useful in string theory.

Question from the audience. If the Laplacian vanishes, is this a standing-wave equation? ^a

Response. No. Both spatial derivatives have the same sign and there is no time derivative. A wave equation requires time dependence with the Lorentzian relative sign. Electrostatics is static precisely because nothing, including a wave, is evolving.

^aLecture timestamp: 00:05:35.

Question from the audience. Is there a curl in two dimensions? ^a

Response. Yes, but it has only one independent component. For $\mathbf{E} = (E_x, E_y)$,

$$\text{curl}_{2d} \mathbf{E} = \partial_y E_x - \partial_x E_y.$$

It may be pictured as the component perpendicular to the board, although an intrinsically two-dimensional world needs no external perpendicular direction. In higher dimensions the corresponding object is naturally an antisymmetric two-form.

^aLecture timestamp: 00:06:06.

8.2 A Derivative with No Preferred Direction

Introduce complex coordinates on the original and image planes:

$$z = x + iy, \quad w = u + iv, \quad w = w(z). \quad (8.5)$$

Here x, y, u, v are real. Locally, the derivative ought to be

$$\frac{dw}{dz} = \lim_{\Delta z \rightarrow 0} \frac{\Delta w}{\Delta z}. \quad (8.6)$$

The complex plane, unlike the real line, permits infinitely many directions of approach. A continuous and ordinarily smooth map need not give the same ratio along each direction.

Compare two simple approaches. Along the x -direction, $dy = 0$, so

$$\left. \frac{dw}{dz} \right|_x = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}. \quad (8.7)$$

Along the y -direction, $dx = 0$, and

$$\left. \frac{dw}{dz} \right|_y = \frac{1}{i} \frac{\partial u}{\partial y} + \frac{\partial v}{\partial y} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y}. \quad (8.8)$$

Equality of real and imaginary parts gives

$$\boxed{\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}} \quad (8.9)$$

These are the Cauchy–Riemann equations. Under the usual differentiability assumptions they are also sufficient for the limit (8.6) to be direction independent. A function with this property is analytic, or holomorphic, at the point.

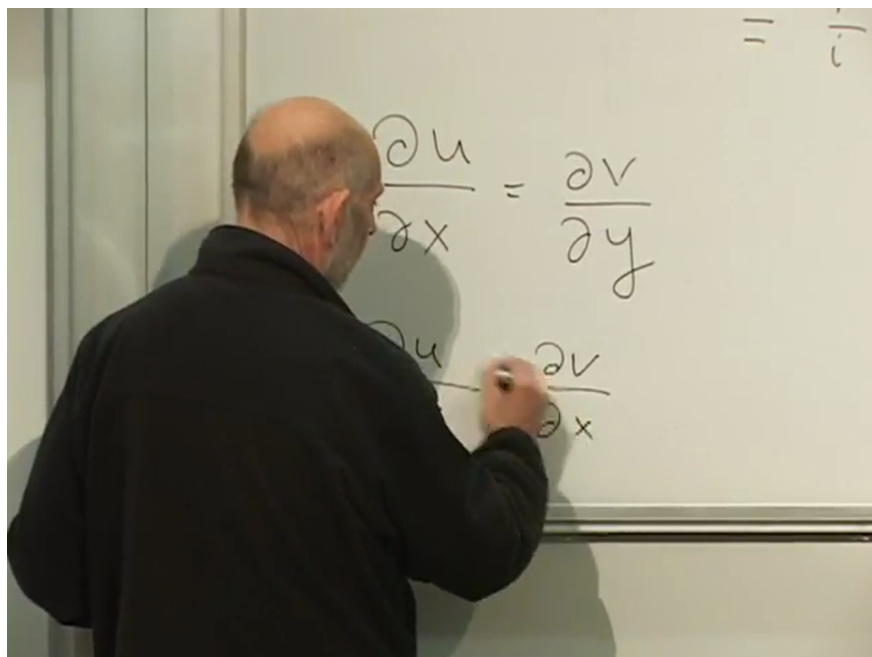


Figure 8.2: The completed Cauchy–Riemann pair, including the relative minus sign.

Lecture frame: 00:19:48.

The lecture briefly distinguishes related terminology. A function holomorphic throughout the finite complex plane is entire. A meromorphic function may have isolated poles. Thus $1/z$ is analytic away from $z = 0$, but it is not entire.

8.3 Harmonic Conjugates and Conformal Geometry

The Cauchy–Riemann equations are coupled first-order equations, but they imply separate second-order equations. Differentiate $\partial_x u = \partial_y v$ with respect to x , differentiate $\partial_y u = -\partial_x v$ with respect to y , and add:

$$\partial_x^2 u = \partial_x \partial_y v, \quad (8.10)$$

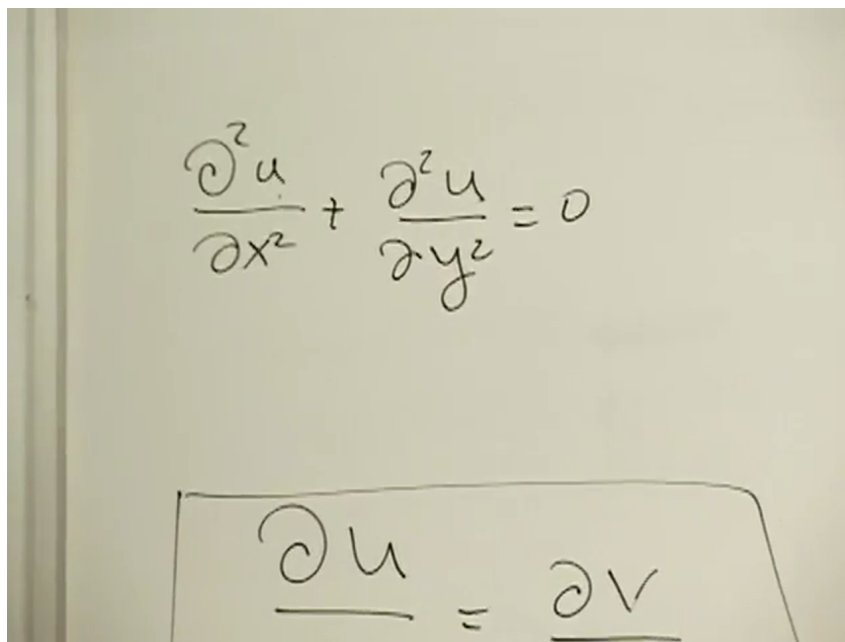
$$\partial_y^2 u = -\partial_y \partial_x v, \quad (8.11)$$

$$\therefore \quad \partial_x^2 u + \partial_y^2 u = 0. \quad (8.12)$$

Repeating the argument with u and v interchanged gives

$$\partial_x^2 v + \partial_y^2 v = 0. \quad (8.13)$$

The real and imaginary parts of an analytic function are therefore harmonic conjugates. Two arbitrary harmonic functions do not automatically make an analytic function; they must still be linked by (8.9).



$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

$$\boxed{\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}}$$

Figure 8.3: The mixed derivatives cancel, leaving Laplace’s equation for one component of the analytic map.

Lecture frame: 00:23:09.

There is also a useful vector interpretation. If $\mathbf{F} = (u, -v)$, then the two Cauchy–Riemann equations state, up to the chosen orientation convention, that both the divergence and the two-dimensional curl of $\mathbf{F}$ vanish. This is the promised return to electrostatics.

8.3.1 Why Angles Are Preserved

Write two small complex displacements in polar form:

$$\delta z = \rho e^{i\theta}, \quad \Delta z = \rho' e^{i\theta'}. \quad (8.14)$$

Their ratio has phase $\theta - \theta'$, exactly the angle between them. Under an analytic map,

$$\frac{\delta w}{\delta z} = \frac{\Delta w}{\Delta z} = w'(z), \quad (8.15)$$

because the derivative does not depend on direction. Rearranging gives

$$\frac{\Delta z}{\delta z} = \frac{\Delta w}{\delta w}. \quad (8.16)$$

The two sides have the same phase, so the angle between the original intervals equals the angle between their images.

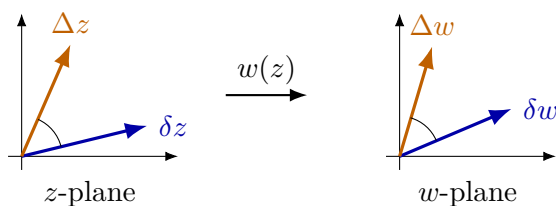
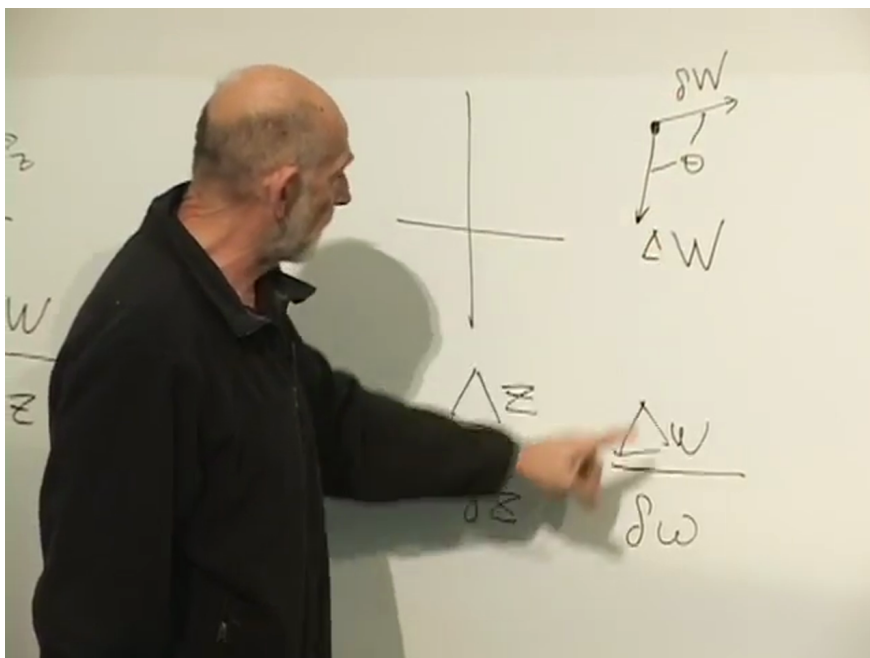


Figure 8.4: The board proof and a clean local reconstruction. Multiplication by $w'(z)$ rotates and rescales every infinitesimal direction by the same amount, preserving their relative angle.

Lecture frame: 00:34:11.

This local argument assumes $w'(z) \neq 0$. At a critical point the map may fail to be locally one-to-one, and one must treat the geometry separately. Away from such points, analytic maps and orientation-preserving conformal maps are locally the same objects.

8.4 Examples that Train the Eye

8.4.1 Squaring and Complex Conjugation

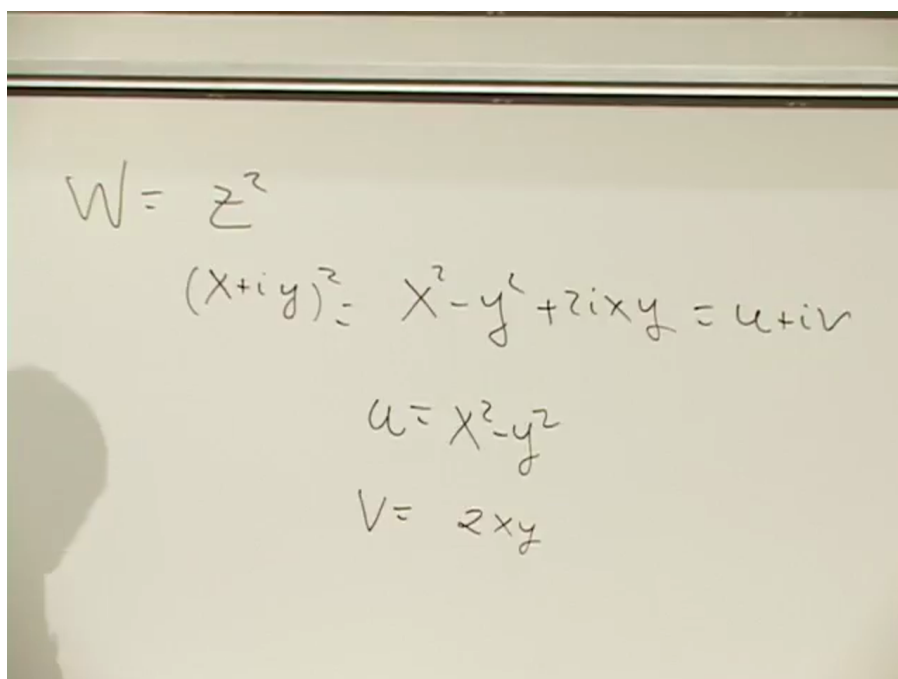
For $w = z^2$,

$$(x + iy)^2 = x^2 - y^2 + 2ixy = u + iv, \quad (8.17)$$

so

$$u = x^2 - y^2, \quad v = 2xy. \quad (8.18)$$

The second derivatives of u are 2 and -2 ; those of v vanish separately. Both are harmonic, and direct differentiation also checks the Cauchy–Riemann equations.



Handwritten notes on a whiteboard showing the squaring map $w = z^2$ and its real and imaginary components u and v .

$$w = z^2$$

$$(x + iy)^2 = x^2 - y^2 + 2ixy = u + iv$$

$$u = x^2 - y^2$$

$$v = 2xy$$

Figure 8.5: The squaring map separated into its real and imaginary harmonic components.

Lecture frame: 00:36:45.

Question from the audience. Does $1/z$ give a nonanalytic example? ^a

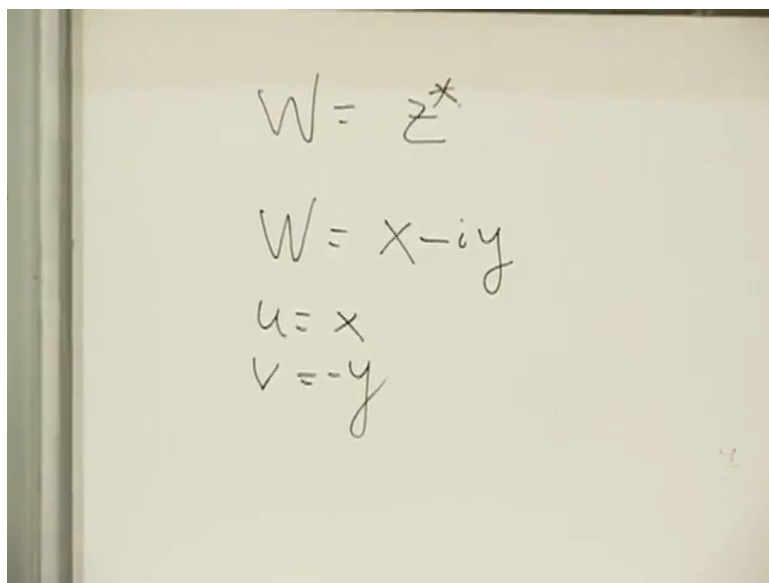
Response. Only at one point. It is analytic wherever $z \neq 0$ and has a pole at the origin. To see a map that fails the Cauchy–Riemann test everywhere, complex conjugation is a better example.

^aLecture timestamp: 00:37:49.

By contrast,

$$w = \bar{z} = x - iy \implies u = x, \quad v = -y \quad (8.19)$$

reverses orientation. Its derivatives fail the first Cauchy–Riemann equation by a sign, so it is not analytic as a function of z .



Handwritten equations on a piece of paper:

$$W = z^*$$
$$W = x - iy$$
$$u = x$$
$$v = -y$$

Figure 8.6: Complex conjugation provides the deliberately nonanalytic counterexample.

Lecture frame: 00:38:38.

8.4.2 The Exponential and Its Inverse

Question from the audience. If e^z is analytic, why should its inverse $\log z$ also be analytic?
^a

Response. Where the inverse exists and $w'(z) \neq 0$, its derivative is $1/w'(z)$. Direction independence is therefore preserved. Globally the logarithm is multivalued, so one chooses a branch and excludes its branch point and cut. The local conformal statement remains valid on that chosen domain.

^aLecture timestamp: 00:40:38.

Euler's formula gives

$$w = e^z = e^{x+iy} = e^x(\cos y + i \sin y), \quad (8.20)$$

hence

$$u = e^x \cos y, \quad v = e^x \sin y. \quad (8.21)$$

For example, $\partial_x u = e^x \cos y = \partial_y v$, and the second Cauchy–Riemann equation follows just as directly.

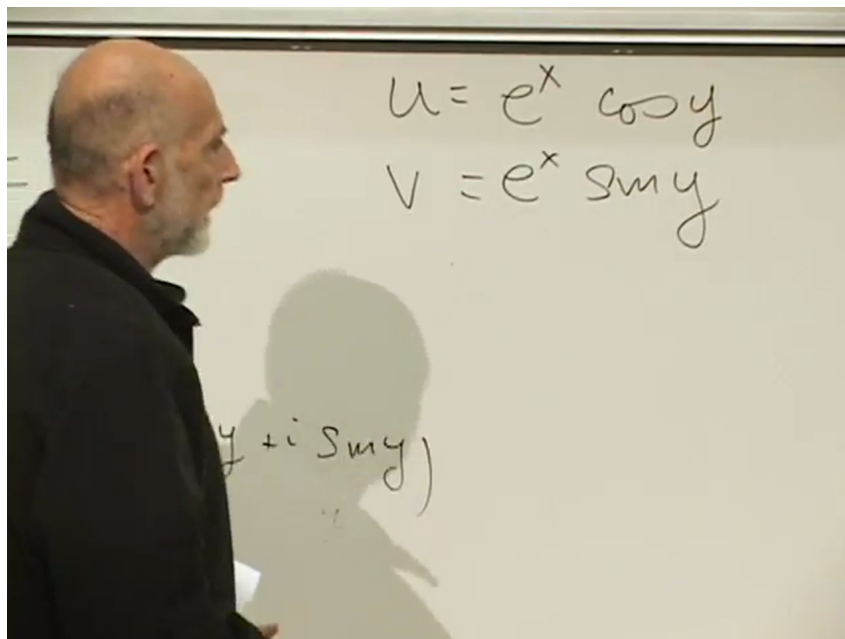


Figure 8.7: The real and imaginary components of the exponential map.

Lecture frame: 00:42:23.

8.5 Half-Plane, Strip, Cylinder, and Disc

8.5.1 The Logarithm Turns Radius into Time

Take the right half-plane, where $-\pi/2 < \theta < \pi/2$, and write

$$z = re^{i\theta}, \quad w = \log z = \log r + i\theta. \quad (8.22)$$

The half-plane maps to the infinite strip

$$-\frac{\pi}{2} < \operatorname{Im} w < \frac{\pi}{2}. \quad (8.23)$$

Rays of fixed θ become horizontal lines, while semicircles of fixed r become vertical lines. The origin is sent to $\operatorname{Re} w = -\infty$, and $r = \infty$ to $\operatorname{Re} w = +\infty$.

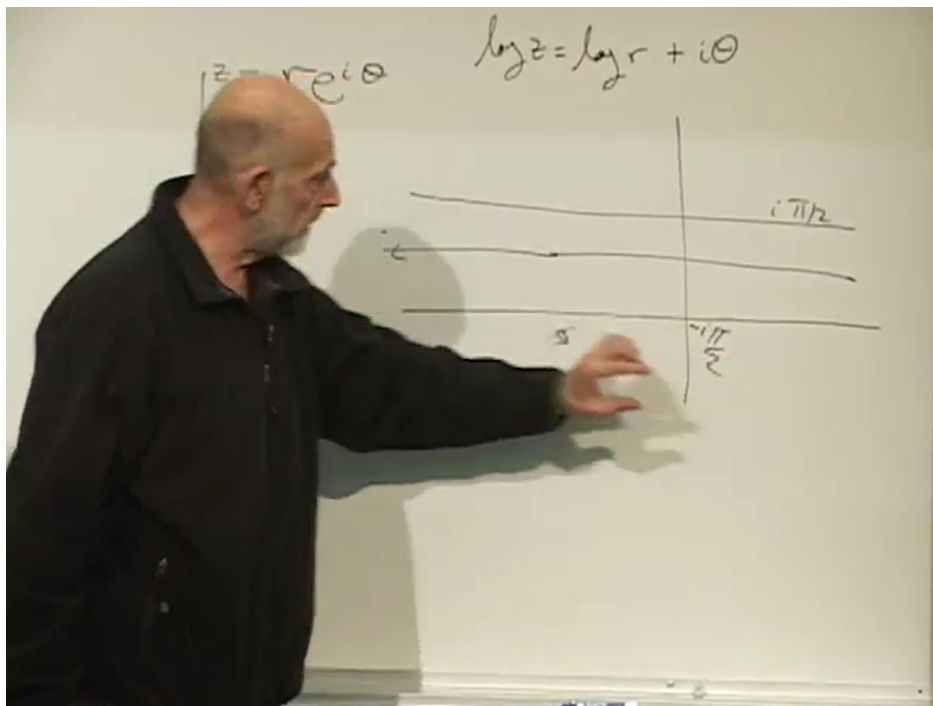


Figure 8.8: The logarithm carries the right half-plane into an infinite strip: polar rays and semicircles become an orthogonal rectangular grid.

Lecture frame: 00:47:48.

The strip is the standard worldsheet of an open string. Its bounded coordinate is the parameter along the string, while translation along the unbounded direction is worldsheet time. Under the inverse exponential map, a time translation becomes a dilation $z \mapsto e^a z$. Thus the same symmetry looks like translation on the strip and uniform expansion on the half-plane.

Question from the audience. Where does a particular value such as $r = 1$ go, and are the coordinates on the strip still polar? ^a

Response. A circle of fixed radius becomes a vertical line at $\text{Re } w = \log r$; in particular $r = 1$ maps to the central line $\text{Re } w = 0$. Polar coordinates describe the z -plane. The w -plane uses ordinary rectangular coordinates $(\log r, \theta)$.

^aLecture timestamp: 00:50:58.

If the full punctured plane is mapped by $\log z$, the angular coordinate has period 2π . The resulting strip therefore has its upper and lower edges identified: it is an infinite cylinder. The half-plane with a boundary describes the open-string worldsheet; the full plane, with no boundary and periodic angle, describes a closed-string cylinder.

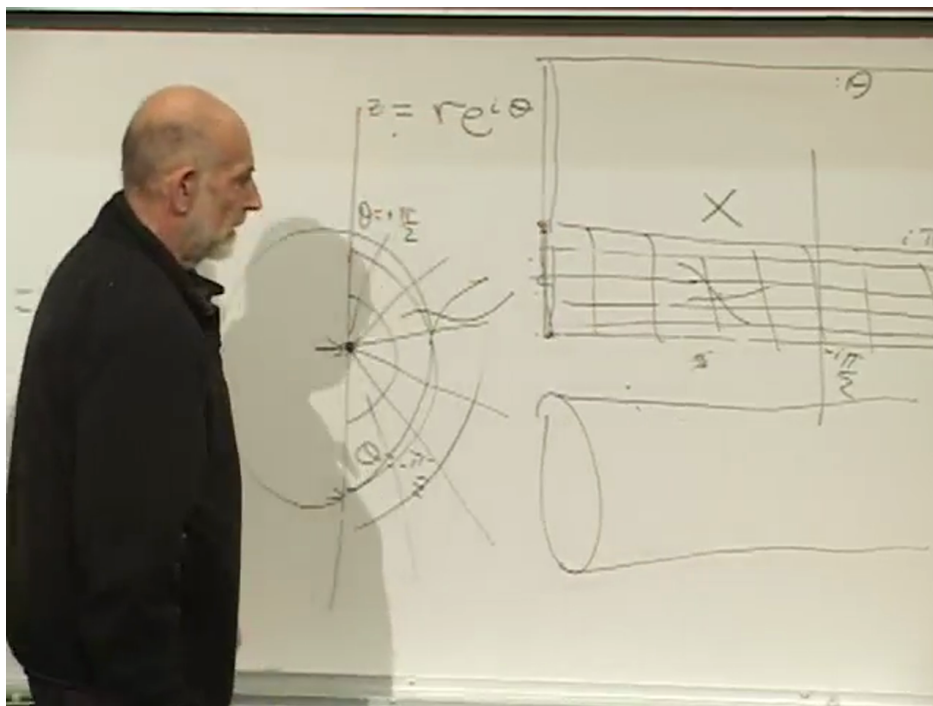


Figure 8.9: The half-plane/strip picture and the periodic identification that turns the full punctured plane into a closed-string cylinder.

Lecture frame: 00:54:23.

8.5.2 A Cayley Map to the Unit Disc

A linear fractional, or Möbius, map carries the right half-plane to the unit disc:

$$w = \frac{z - 1}{z + 1}. \quad (8.24)$$

It sends $z = 0$ to -1 , $z = \infty$ to $+1$, and the imaginary axis to $|w| = 1$. If $z = ir$,

$$\left| \frac{ir - 1}{ir + 1} \right| = 1. \quad (8.25)$$

Points with $\text{Re } z > 0$ lie inside the circle; the left half-plane maps outside it. ²

²Editorial clarification: During the live derivation the reciprocal $(z + 1)/(z - 1)$ is written while the interior-disc image is described. That reciprocal has the same unit-circle boundary but sends the right half-plane to the exterior. Equation (8.24) is the interior-disc convention consistent with the stated geometry.

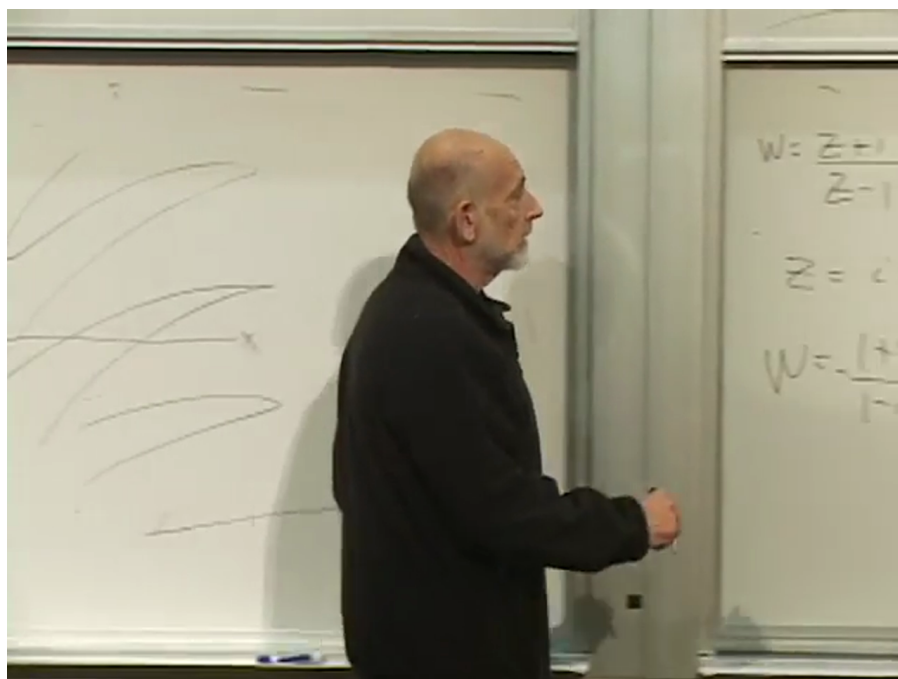


Figure 8.10: The linear-fractional construction of the unit-circle boundary. The reciprocal convention visible on the board exchanges interior and exterior.

Lecture frame: 00:58:38.

Möbius maps send the generalized collection of circles and straight lines to circles and straight lines. The radial and circular grid on the half-plane consequently becomes two orthogonal families of arcs on the disc. A straight diameter is simply the limiting case of a circle whose center has gone to infinity.

Question from the audience. Are the arcs drawn inside the disc really circles? ^a

Response. Yes. Their centers occur at different locations, often outside the visible disc. The diameter is the infinite-radius limit of the same family.

^aLecture timestamp: 01:00:55.

Question from the audience. Does the disc drawing mean that it is a sphere, and where does the other half-plane go? ^a

Response. The drawing here is a flat disc, not a sphere. The right half-plane maps inside its boundary and the left half-plane maps outside. If one adds the point at infinity to the entire complex plane, that extended plane can be represented as a sphere by stereographic projection; that is a different global statement.

^aLecture timestamp: 01:01:21.

Question from the audience. Which disc curves correspond to the original radial spokes and constant-radius semicircles? ^a

Response. The positive real axis becomes the diameter joining -1 and $+1$. Other rays become circular arcs joining those same boundary points. Semicircles of fixed radius become the orthogonal family that begins and ends on the unit circle.

^aLecture timestamp: 01:02:59.

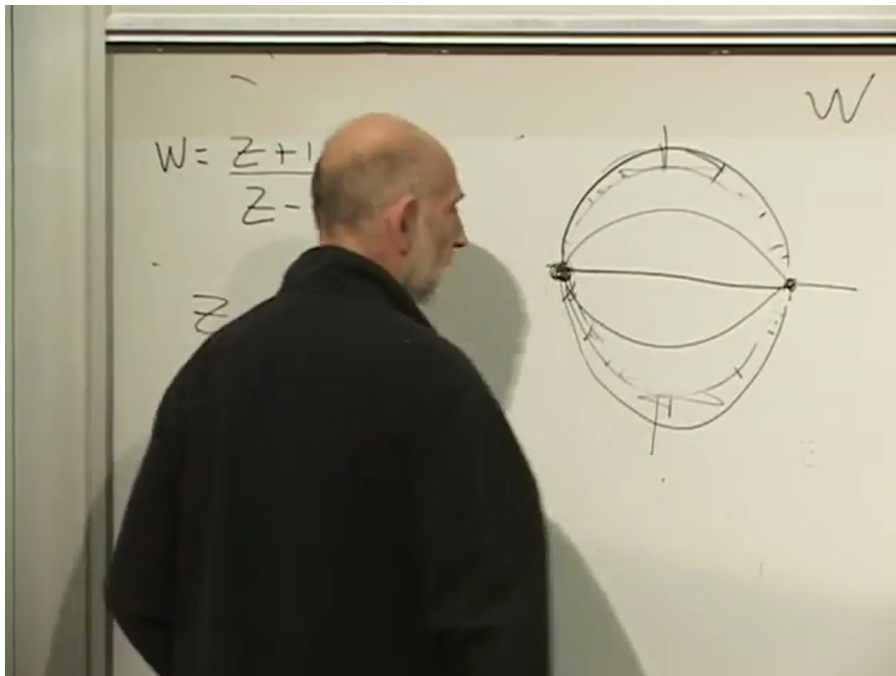


Figure 8.11: Lines and semicircles in the half-plane become orthogonal circle-and-line families in the disc.

Lecture frame: 01:03:48.

8.6 The Four-String Worldsheet

We can now return to scattering. Two open strings enter, join into one string, propagate for an arbitrary interval, and split into two outgoing strings. In strip coordinates the worldsheet looks like a long ribbon slit from each end—the "sports bandage" picture. The intermediate lifetime is not chosen in advance. It is one of the histories over which the amplitude is integrated.

Let (σ, τ) be coordinates on the worldsheet and let

$$X^\mu = X^\mu(\sigma, \tau) \quad (8.26)$$

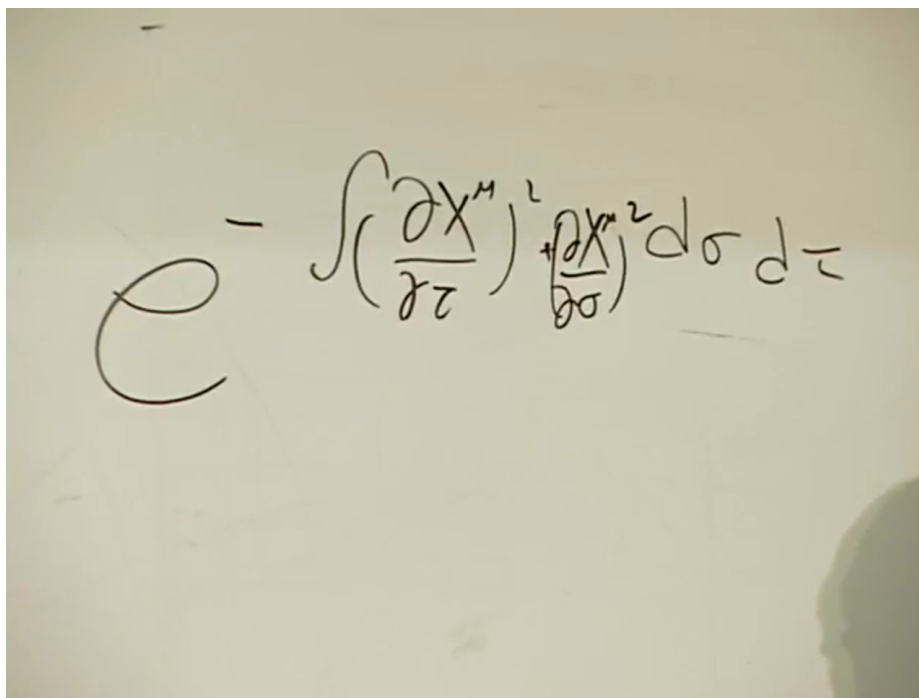
give its position in spacetime. The labels σ, τ live on the two-dimensional parameter surface; the index μ labels spacetime coordinates. In the bosonic theory $\mu = 0, \dots, 25$; in the superstring case there are ten spacetime coordinates.

After Wick rotation, the free embedding fields have Euclidean action

$$S_E[X] = \frac{1}{4\pi\alpha'} \int_\Sigma d\sigma d\tau [\partial_\tau X^\mu \partial_\tau X_\mu + \partial_\sigma X^\mu \partial_\sigma X_\mu]. \quad (8.27)$$

The corresponding path integral sums over every embedding of the chosen worldsheet into spacetime:

$$\int \mathcal{D}X e^{-S_E[X]}. \quad (8.28)$$



The image shows a handwritten mathematical expression on a piece of paper. The expression is $e^{-\int (\frac{\partial X^\mu}{\partial \tau})^2 + (\frac{\partial X^\mu}{\partial \sigma})^2 d\sigma d\tau}$. The 'e' is written as a large circle. The integral is followed by a minus sign and the integrand in parentheses. The integrand consists of two terms: $(\frac{\partial X^\mu}{\partial \tau})^2$ and $(\frac{\partial X^\mu}{\partial \sigma})^2$, separated by a plus sign. The differentials $d\sigma d\tau$ are at the end of the integral.

Figure 8.12: The completed positive Euclidean worldsheet weight. The two derivative terms are summed over all spacetime components.

Lecture frame: 01:11:34.

The slit strip has one connected boundary. Conformal freedom therefore allows it to be represented as a disc. Four remote ends of the strips become four marked points on the boundary. Deforming the marked points changes the apparent separation between joining and splitting in strip coordinates, but it does not change the underlying surface or physics.

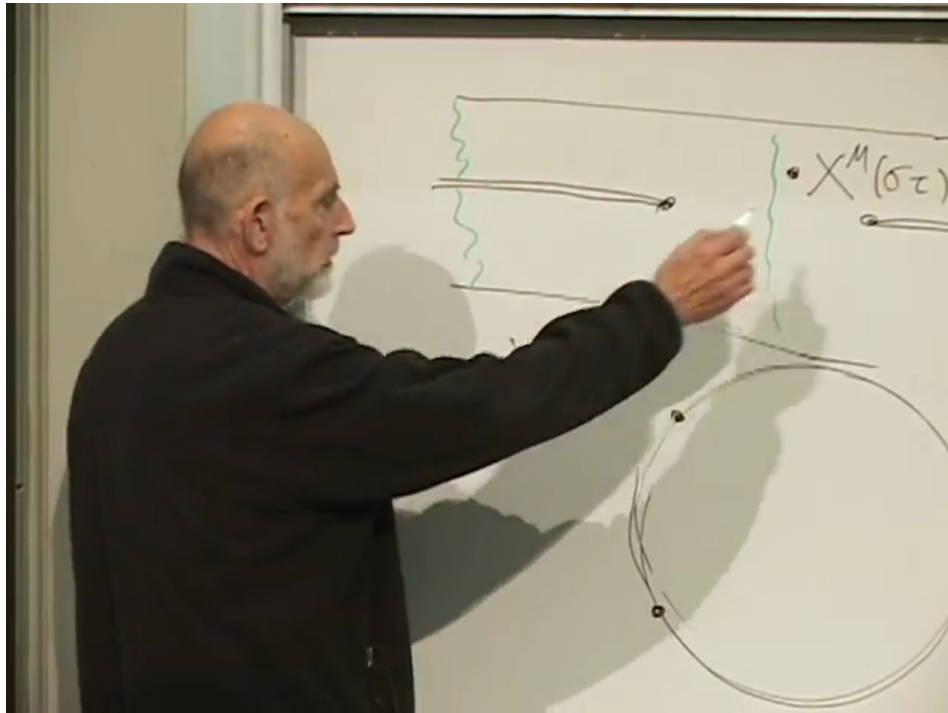


Figure 8.13: The slit-strip worldsheet and its disc representation with boundary points for the external strings.

Lecture frame: 01:15:18.

The detailed grid lines used to picture this conformal map carry no physical content. What matters is the topology, the boundary, and the marked external points. The same worldsheet can be stretched into different channel pictures without introducing a new interaction.

8.7 External States as Boundary Insertions

The bare integral (8.28) does not yet say which momenta enter and leave. For a scalar open-string state of momentum k_i^μ , insert on the boundary the vertex operator

$$V_i(z_i) = e^{ik_i \cdot X(z_i)}, \quad k_i \cdot X = k_{i\mu} X^\mu. \quad (8.29)$$

The point z_i is a position on the worldsheet boundary. It is not a spacetime coordinate and not a direction of momentum.

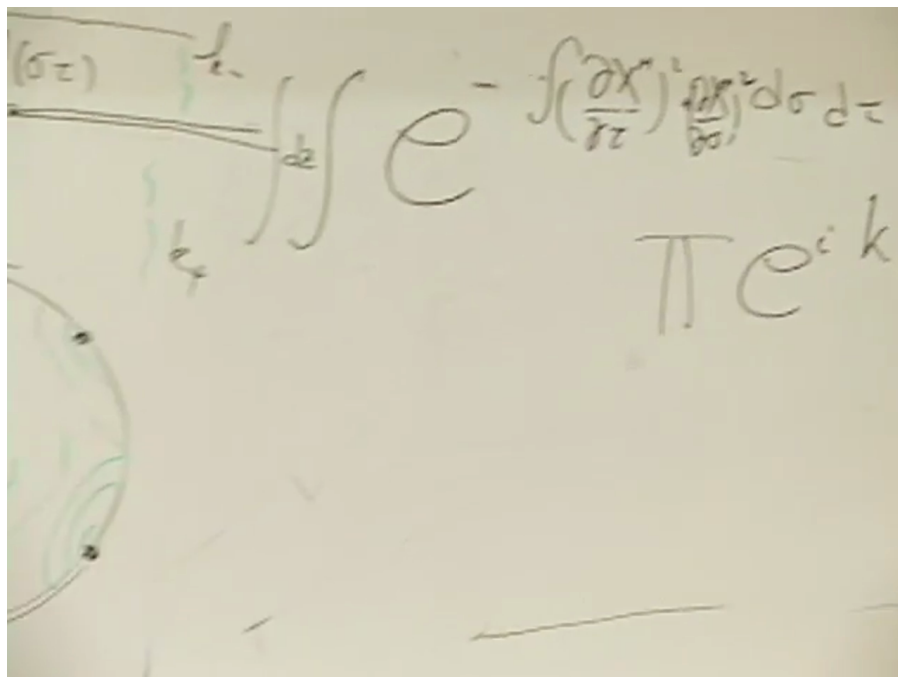


Figure 8.14: The bulk Euclidean weight and the product of boundary momentum insertions are displayed as distinct factors.

Lecture frame: 01:18:58.

For n external open strings, the schematic prescription is

$$\begin{aligned} \mathcal{A}_n(k_1, \dots, k_n) = & \int_{\mathcal{M}_{0,n}^\partial} d\mu(z_i) \int \mathcal{D}X \\ & \times e^{-S_E[X]} \prod_{i=1}^n e^{ik_i \cdot X(z_i)}. \end{aligned} \quad (8.30)$$

The functional integral is over embeddings X^μ . The finite-dimensional integral is over inequivalent boundary insertion positions. Polarization factors and the gauge-fixing measure are suppressed at this stage, as they are in the lecture. ³

Question from the audience. If the external momenta are fixed, why are they integrated over? ^a

Response. They are not. The k_i are fixed labels of the scattering amplitude. What is integrated are the embedding fields and the worldsheet locations z_i at which those fixed momenta are inserted. The latter integral replaces the integral over the intermediate lifetime in the strip picture.

^aLecture timestamp: 01:20:42.

³Editorial clarification: A complete gauge-fixed string amplitude also contains the appropriate ghost contribution, state-dependent vertex factors, coupling constants, and normalization. Equation (8.30) records the embedding-field and position dependence developed in this lecture.

Question from the audience. Is the product of vertex factors inside the path integral, and is it part of the action? ^a

Response. It multiplies e^{-S_E} inside the functional integral. One may equivalently exponentiate the factors and regard them as localized source terms added to the action, but they are conceptually vertex operators rather than part of the free bulk action.

^aLecture timestamp: 01:21:39.

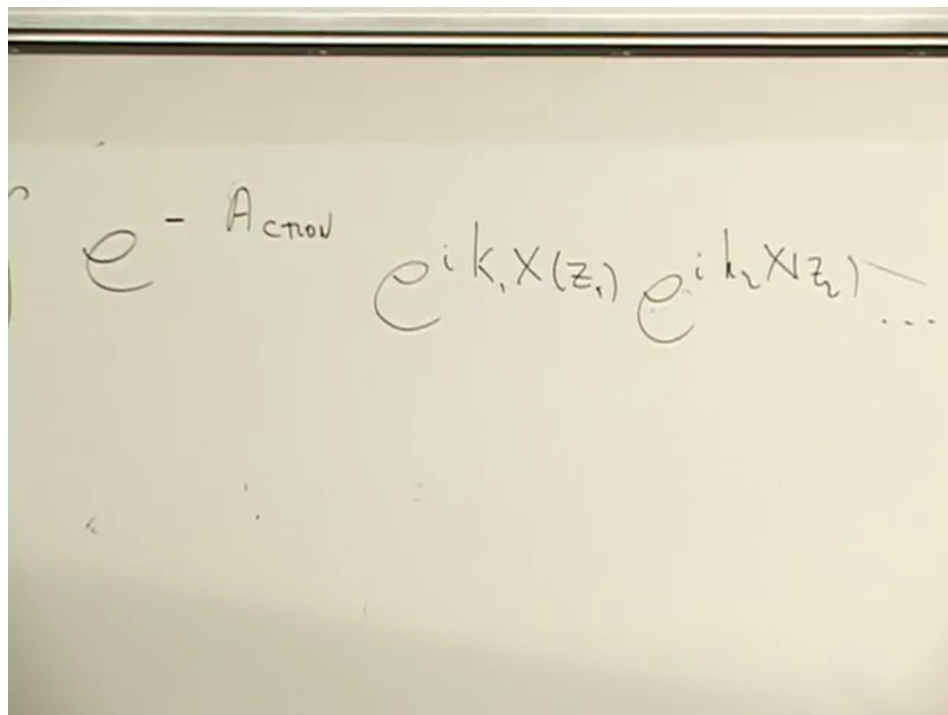


Figure 8.15: The path integral written with separate factors $e^{ik_i \cdot X(z_i)}$ for successive external particles.

Lecture frame: 01:23:23.

Question from the audience. Why does $e^{ik \cdot X}$ still contain i after the worldsheet Wick rotation? ^a

Response. The rotation changes the worldsheet time and converts the bulk weight from e^{iS} to e^{-S_E} . The i in a momentum eigenfunction is the Fourier phase that labels spacetime momentum. It has a different role and remains.

^aLecture timestamp: 01:23:42.

Question from the audience. Does the position z_i determine the direction of the momentum vector? ^a

Response. No. Direction is encoded in the spacetime components k_i^μ . The insertion point z_i says where that momentum enters the worldsheet description; $X^\mu(z_i)$ is among the field values summed over by $\mathcal{D}X$.

^aLecture timestamp: 01:24:19.

Question from the audience. Do the drawn angles of the bandage legs encode the directions of the physical momenta? ^a

Response. No. The bandage is parameter space, not an image of the string in physical spacetime. Its legs may be drawn horizontally, tilted, or conformally bent. Physical directions are carried by the components k^μ , while $X^\mu(\sigma, \tau)$ is integrated over.

^aLecture timestamp: 01:25:07.

8.8 The Gaussian Integral as Electrostatics

The action (8.27) is quadratic in every component X^μ , and the insertions are linear sources once their exponentials are combined with the action. The functional integral is therefore Gaussian. At its saddle point, each spacetime component obeys a Poisson equation of the schematic form

$$\nabla^2 X^\mu(z) \propto i\alpha' \sum_i k_i^\mu \delta(z - z_i), \quad (8.31)$$

with the open-string boundary condition on the rim of the disc. Each X^μ behaves like an independent electrostatic potential, and each momentum component k_i^μ behaves like the corresponding charge placed at z_i .

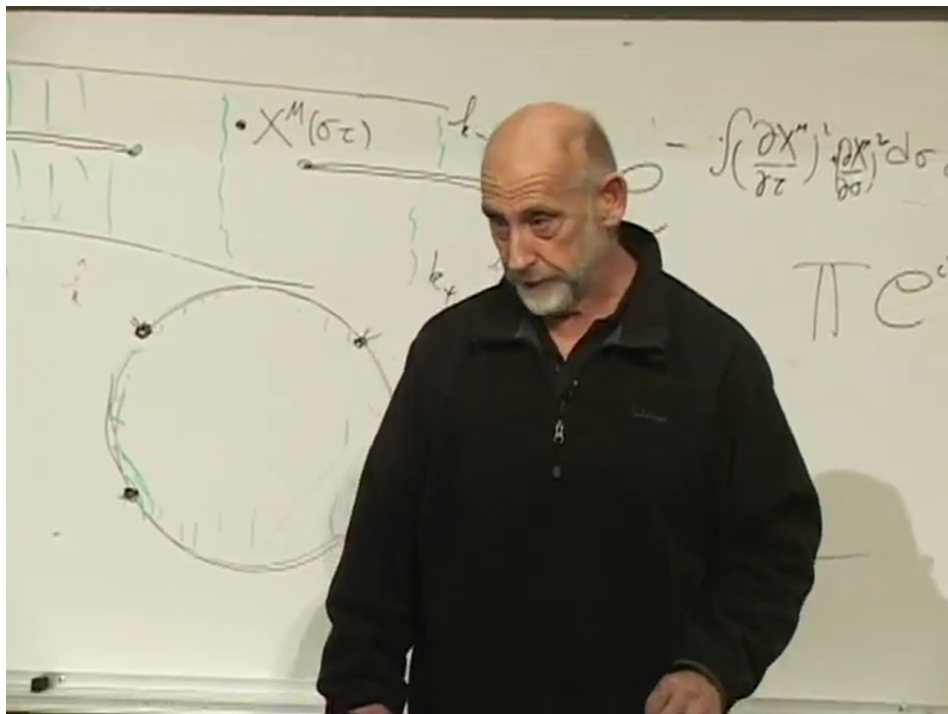


Figure 8.16: The disc with boundary sources is interpreted as one independent two-dimensional electrostatic problem for each spacetime component.

Lecture frame: 01:28:58.

Solve the electrostatic problem, evaluate its energy, exponentiate the negative energy, and finally integrate over the inequivalent charge positions. In standard open-string normalization, the

embedding-field correlator supplies the Koba–Nielsen factor

$$\left\langle \prod_i e^{ik_i \cdot X(z_i)} \right\rangle \propto (2\pi)^D \delta^{(D)} \left(\sum_i k_i \right) \prod_{i < j} |z_i - z_j|^{2\alpha' k_i \cdot k_j}. \quad (8.32)$$

The delta function expresses momentum conservation. The product contains the electrostatic pair energies. For the four scalar boundary insertions used here, fixing three positions turns the remaining integral into the Euler beta integral encountered in the preceding lectures.

Question from the audience. Are there twenty-six electrostatic energies or only one? ^a

Response. There is one independent potential and field for each spacetime component. Their quadratic energies have no cross terms in an orthonormal target-space basis, and the contributions are added to one total action before exponentiation.

^aLecture timestamp: 01:30:53.

Question from the audience. Is the angular coordinate of the disc the physical scattering angle? ^a

Response. No. It is a worldsheet position parameter. Its inequivalent combination is related to the intermediate propagation interval and is integrated. The physical scattering angle is encoded through Lorentz invariants built from the fixed momenta.

^aLecture timestamp: 01:32:10.

Question from the audience. Why does this construction fail away from the critical dimension? ^a

Response. The formal Gaussian integrals alone are not the whole consistency condition. Gauge fixing introduces a conformal anomaly whose matter and ghost contributions cancel only at the critical dimension: $D = 26$ for the bosonic string and $D = 10$ for the superstring. Without that cancellation the supposed conformal symmetry of the quantum worldsheet is lost.

^aLecture timestamp: 01:32:35.

Question from the audience. Have we already assumed three large spatial dimensions and compact extra dimensions? ^a

Response. No. Compactification has not yet entered the construction. At this stage the bosonic coordinates still range over all twenty-six target-space directions, or ten in the superstring case. How a subset can remain macroscopically large is the subject of the next discussion.

^aLecture timestamp: 01:33:38.

Question from the audience. Is the connection with electrostatics simply that the coordinates satisfy Laplace’s equation? ^a

Response. Exactly. Away from the insertion points each X^μ is harmonic; the vertex operators supply delta-function sources. That is mathematically the same boundary-value problem as a static two-dimensional potential with point charges, whether or not we use the electrostatic vocabulary.

^aLecture timestamp: 01:33:57.

8.8.1 Conformal Redundancy and the Number of Moduli

The disc has a three-real-parameter conformal automorphism group, $\text{PSL}(2, \mathbb{R})$. Three ordered boundary points may therefore be fixed at convenient locations. For four external strings only the fourth position remains to be integrated; for n ordered boundary insertions there are

$$n - 3 \quad (8.33)$$

real position integrals at tree level.

Question from the audience. Are there four independent angular positions around the disc?

^a

Response. Not physically. One initially writes four positions, but disc conformal symmetry can place any three wherever convenient. Only the relative location of the fourth remains for the two-in/two-out amplitude.

^aLecture timestamp: 01:34:37.

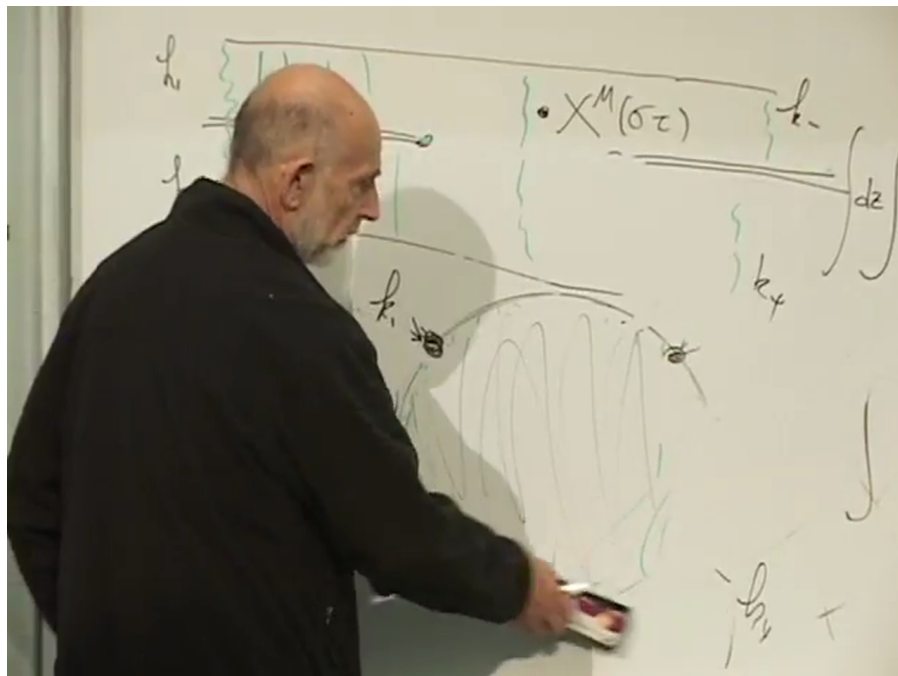


Figure 8.17: Four boundary insertions on the disc. Conformal freedom fixes any three, leaving one four-point modulus.

Lecture frame: 01:35:18.

Question from the audience. What if the two split points do not occur at the same height in the strip drawing? ^a

Response. A conformal map can move that apparent mismatch away. Different drawings are coordinate representatives of the same modulus. The invariant integration is the one over the remaining boundary position, equivalently over the propagation interval.

^aLecture timestamp: 01:35:24.

Question from the audience. For more external particles, how many position parameters remain? ^a

Response. Fix three insertion points with conformal symmetry and integrate the other $n - 3$. This is the boundary-moduli count for an ordered tree-level open-string amplitude.

^aLecture timestamp: 01:36:00.

Question from the audience. How is this related to M-theory? ^a

Response. Not yet. The construction here is perturbative open-string scattering on a worldsheet. Its relation to M-theory requires later duality and nonperturbative ideas, so the question is explicitly deferred.

^aLecture timestamp: 01:36:37.

Question from the audience. Is a different vertical or crossed strip picture a different calculation? ^a

Response. No. One may hold a convenient point fixed and let the remaining modulus range through its full interval. The strip pictures and the boundary-point integral on the disc are different slicings of the same integral, which is why all external strings can be treated symmetrically.

^aLecture timestamp: 01:36:49.

8.9 What the Prescription Does Not Yet Solve

The disc formula extends immediately to more external strings: add more vertex operators and integrate their $n - 3$ inequivalent boundary positions. It is a precise tree-level prescription for a consistent string model, but it is not yet a realistic model of observed particle physics. The simple bosonic theory has the wrong spacetime dimension and additional physical defects. Superstrings improve the spectrum and lower the critical dimension to ten, but six spatial dimensions still have to be compactified.

Question from the audience. Does this simple model already agree with particle collision experiments? ^a

Response. No. The resonance structure historically resembles hadronic scattering, but the twenty-six-dimensional bosonic model cannot be a direct description of observed particles. The formalism defines scattering amplitudes for the model; matching nature requires a realistic spectrum, compactification, interactions, and vacuum choice.

^aLecture timestamp: 01:39:06.

Question from the audience. Will compactification reduce the extra dimensions and produce realistic particle data? ^a

Response. It can reduce the effective macroscopic dimension and generate particle spectra and gauge groups. It is also where much of the real mathematical difficulty enters. Many consistent compactifications are known, and selecting a unique model of observed elementary particles remains unresolved.

^aLecture timestamp: 01:41:10.

The lecture emphasizes this distinction rather than advertising a finished theory. String constructions are mathematically definite and have produced deep information about gravity and geometry, yet their multiplicity is itself part of the physical problem.

Question from the audience. Is there a short way to remember why the bosonic critical dimension is twenty-six? ^a

Response. The earlier oscillator calculation used the regularized zero-point sum $1+2+3+\cdots = -1/12$. The transverse oscillator contribution involves twenty-four directions, and restoring the two longitudinal spacetime coordinates gives $24+2=26$. This is a mnemonic for the earlier consistency argument, not a new derivation here.

^aLecture timestamp: 01:43:19.

Question from the audience. Are the inserted momenta k_i themselves quantized? ^a

Response. In noncompact spacetime their components are continuous. What is quantized is the oscillator level and hence the allowed mass-squared spectrum. Compact dimensions may quantize the corresponding momentum and winding components, but that belongs to the next discussion.

^aLecture timestamp: 01:43:46.

Chapter 9

Strings, Gravity, Compactification, and T-Duality

Overview. A point particle can move consistently on essentially any smooth geometry, but a quantum string cannot. Its infinitely many zero-point oscillations probe the target space at successively shorter scales, and the effective geometry changes as those oscillations are restored. Requiring that this process have a fixed point gives $R_{\mu\nu} = 0$, the vacuum Einstein equation. We then turn from this consistency condition to compactification. A circle supplies both Kaluza–Klein momentum states and string winding states; exchanging them identifies a circle of radius R with one of radius α'/R .

9.1 Consistency as a Selection Principle

There are two questions to answer. Why does gravity occur in string theory, and why is string theory far less permissive than point-particle mechanics? They are really one question. A background geometry is acceptable only if a string can propagate on it without retaining an arbitrary dependence on how finely we resolve the string.

Begin with the control problem: a point particle constrained to a smooth curved surface. Classically it follows a geodesic, the intrinsic analogue of a straight line. One can approximate the motion by many short segments, write a difference equation, and then let the segment length tend to zero. The result is an ordinary differential equation with a well-defined continuum limit.

Quantum mechanics changes the sum over histories, not the conclusion. Instead of selecting one classical path, we sum amplitudes over paths on the surface. Time may again be divided into small steps, and the limit is equivalently described by a Schrödinger equation on the curved configuration space. For a smooth surface, the limit settles to a definite propagator and spectrum.

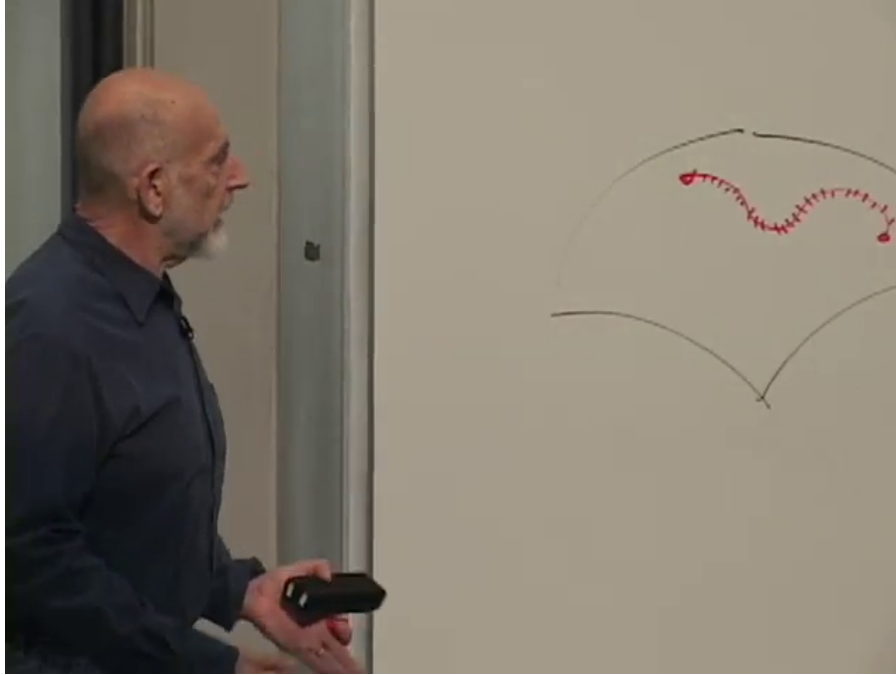


Figure 9.1: A smooth path on the surface and a jagged path from the time-sliced quantum sum. Both remain intrinsic to the target surface.

Lecture frame: 00:05:10.

Question from the audience. In the path integral, are all of the paths required to remain inside the surface? ^a

Response. Yes. The surface is the particle's configuration space, not a sheet embedded in a larger room through which the particle may take a shortcut. Every intermediate point belongs to the surface, and all lengths and actions are computed with its intrinsic metric.

^aLecture timestamp: 00:05:39.

The corresponding question for a string is harder. A string is not specified only by a center-of-mass position. It carries an infinite tower of oscillators, so improving the resolution means admitting more dynamical degrees of freedom. The theory is consistent only if observables approach a limit as that tower is restored.

9.2 The Point Particle on a Sphere

The sphere is the simplest curved test geometry. Let its radius be R , and first place a particle of mass m on it. For motion along a great circle,

$$E = \frac{1}{2}mv^2 = \frac{p^2}{2m}, \quad L = pR. \quad (9.1)$$

Consequently,

$$E = \frac{L^2}{2mR^2} = \frac{L^2}{2I}, \quad I = mR^2. \quad (9.2)$$

The radius enters through the moment of inertia. Quantizing the angular momentum produces the familiar rigid-rotor spectrum, and nothing depends on an ultraviolet cutoff.

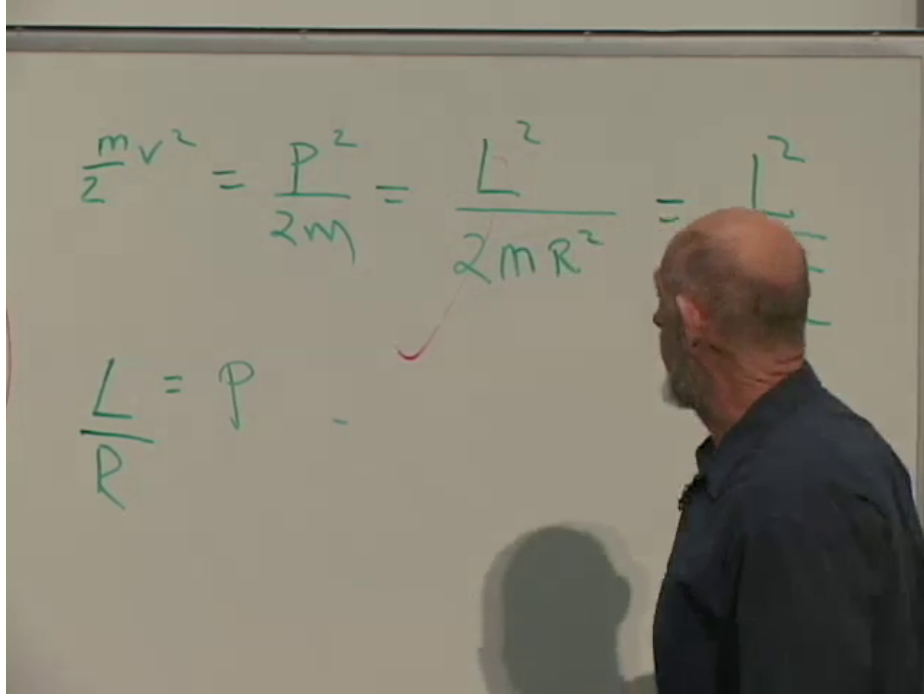


Figure 9.2: The point-particle control calculation: $p^2/(2m) = L^2/(2mR^2) = L^2/(2I)$.

Lecture frame: 00:11:25.

This untroubled example is important. Curvature by itself is not the problem. The problem appears when curvature is coupled to the internal quantum extension of a string.

9.3 The Vacuum String Is Not Pointlike

Consider one transverse coordinate of an open string and subtract its center of mass. The oscillator expansion used here is schematic:

$$X(\sigma) - X_{\text{cm}} \sim \sum_{n=1}^{\infty} \frac{a_n + a_n^\dagger}{\sqrt{n}} \cos(n\sigma). \quad (9.3)$$

The essential ingredients are the cosine modes, the oscillator operators, and the factor $1/\sqrt{n}$. Exact normalizations play no role in the size argument. ¹

¹Editorial clarification: The precise expansion contains convention-dependent factors of α' , numerical constants, and a choice of oscillator notation. Equation (9.3) retains only the dependence needed for the ultraviolet estimate, just as in the board derivation.

The image shows two equations written on a chalkboard. The top equation is the structural oscillator expansion:
$$X(\sigma) = \frac{1}{\alpha'} \sum_n \frac{(a_n^\dagger + a_n^-)}{n} \cos(n\sigma)$$
The bottom equation is the double sum obtained by squaring the top equation:
$$\sum_{n,m} \frac{(a_n^\dagger + a_n^-)(a_m^\dagger + a_m^-)}{\sqrt{n}\sqrt{m}} \cos(n\sigma) \cos(m\sigma)$$

Figure 9.3: The structural oscillator expansion and the double sum obtained on squaring it.

Lecture frame: 00:15:20.

The squared distance from the center of mass in the oscillator vacuum is

$$\begin{aligned} \langle (X(\sigma) - X_{\text{cm}})^2 \rangle_0 &\sim \sum_{n,m \geq 1} \frac{\cos(n\sigma) \cos(m\sigma)}{\sqrt{nm}} \\ &\quad \times \langle 0 | (a_n + a_n^\dagger) \\ &\quad \times (a_m + a_m^\dagger) | 0 \rangle. \end{aligned} \quad (9.4)$$

Most terms vanish. Two annihilation operators kill the ket; two creation operators leave a state orthogonal to the bra; and a creation operator followed by an annihilation operator also has zero vacuum expectation in this ordering. The surviving contraction is

$$\langle 0 | a_n a_m^\dagger | 0 \rangle = \delta_{nm}. \quad (9.5)$$

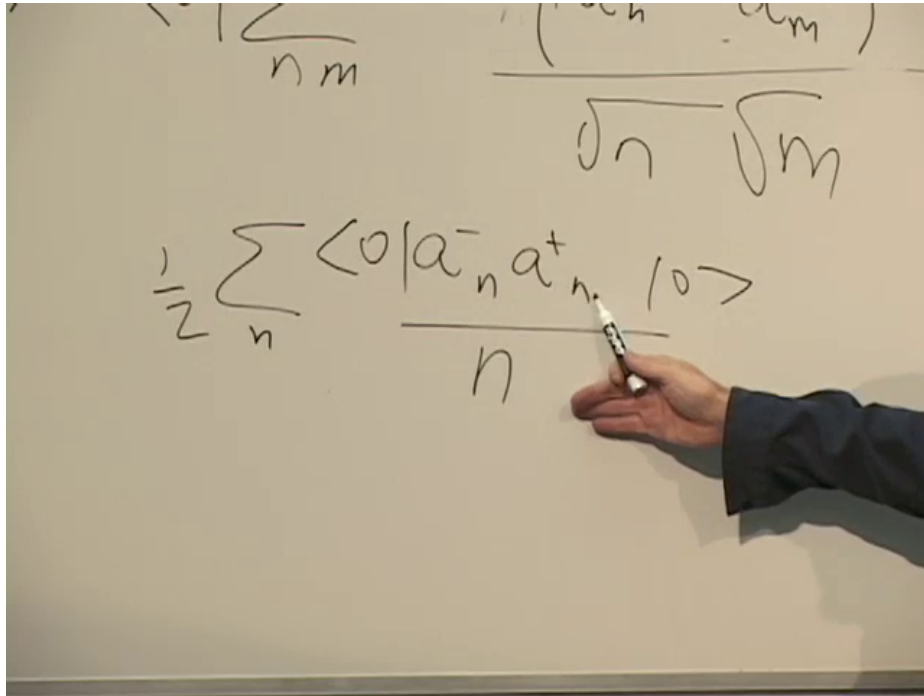


Figure 9.4: The surviving vacuum contraction collapses the double oscillator sum to a single sum.

Lecture frame: 00:20:10.

Thus

$$\langle (X(\sigma) - X_{\text{cm}})^2 \rangle_0 \sim \sum_{n=1}^{\infty} \frac{\cos^2(n\sigma)}{n}. \quad (9.6)$$

At a generic point along the string, or after averaging over σ , $\cos^2(n\sigma)$ contributes roughly $1/2$. With a mode cutoff N ,

$$\langle (X - X_{\text{cm}})^2 \rangle_{0,N} \sim \frac{1}{2} \sum_{n=1}^N \frac{1}{n} = \frac{1}{2} \log N + O(1). \quad (9.7)$$

The image shows a chalkboard with three equations written in chalk. The top equation is $\langle 0 | X | 0 \rangle = \langle 0 | \sum_{n,m} \frac{(\pm a_n^-}{\sqrt{n}}$. The middle equation is $\frac{1}{2} \sum_n \frac{\langle 0 | a_n^- a_n^+ | 0 \rangle}{n}$. The bottom equation is $\langle X^2 \rangle = \frac{1}{2} \sum_n \frac{1}{n}$.

Figure 9.5: The board arrives at the harmonic sum controlling the ground-state spread.

Lecture frame: 00:22:05.

Keeping only, say, the first twenty-five modes gives a finite fuzzy object, but it is not yet the continuum string. Removing the cutoff causes the mean-square spread to grow logarithmically. This conclusion should not be misread as a proof that flat-space string theory is inconsistent. In flat space the center-of-mass and oscillator motions separate, and delicate cancellations in physical amplitudes preserve the expected locality. The growing size is a warning. Curvature tells us whether the warning becomes a failure.

9.4 When Internal Spreading Changes the Geometry

Place the string near the north pole of the sphere. With every oscillator turned off, it behaves like the point particle just studied. As higher modes are admitted, the string becomes a fluctuating tangle spread over the cap. Its center may still travel around the sphere, but the object whose moment of inertia enters the rotational energy is now extended.

For an axis through the center of the sphere and the poles,

$$I = \int r_{\perp}^2 dm, \quad (9.8)$$

where $r_{\perp}$ is the perpendicular distance from a string element to that axis. A point at the equator of the orbit has the largest relevant lever arm. Portions of a string that droop toward the axis contribute less. At fixed total mass this lowers the effective moment of inertia:

$$I_{\text{eff}} < mR^2. \quad (9.9)$$

The rotational energy therefore looks like that of a point object on a smaller sphere,

$$E_{\text{rot}} = \frac{L^2}{2I_{\text{eff}}} \equiv \frac{L^2}{2mR_{\text{eff}}^2}, \quad R_{\text{eff}} < R. \quad (9.10)$$

Question from the audience. Why should the moment of inertia tend toward zero if the string is spreading over a larger region? ^a

Response. The statement is about the effective description under successive coarse-graining, not a claim that the raw extended distribution immediately has zero inertia. The first band of modes turns the point into a fuzzy lump that moves as though it lived on a smaller sphere. The next band resolves structure inside each such lump and repeats the same renormalization. Iterating the comparison drives the effective radius inward.

^aLecture timestamp: 00:33:23.

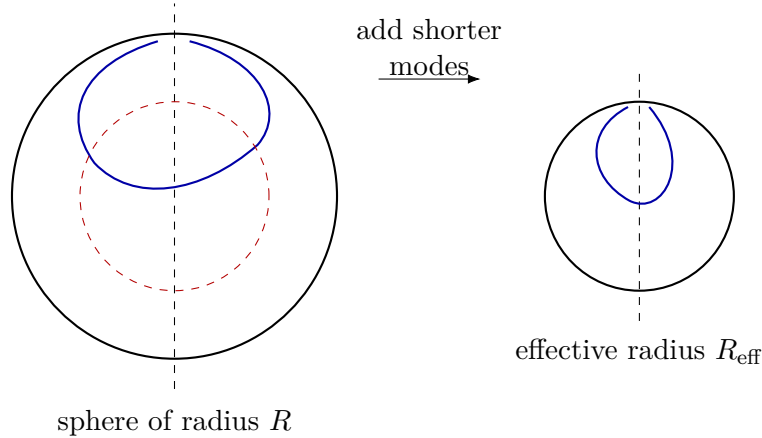
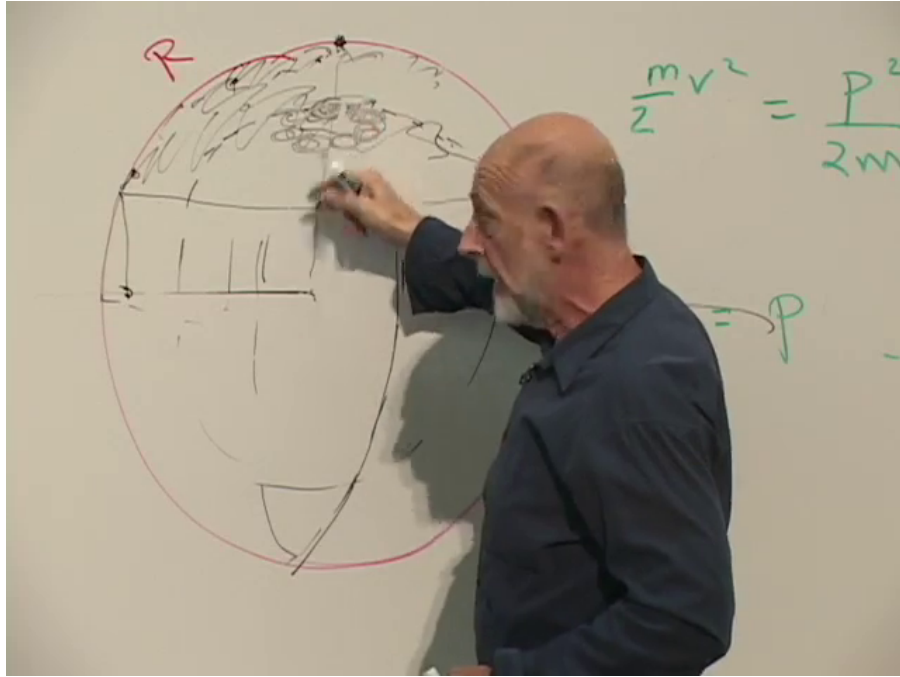


Figure 9.6: The north-pole string on the board and a clean reconstruction of the effective-radius argument. The dashed inner circle is an effective geometry, not a second physical shell inside the sphere.

Lecture frame: 00:35:15.

This is a renormalization-group statement. A finite cutoff makes the string a lump with finite internal resolution. Raising the cutoff changes the mechanics of that lump. On a positively curved sphere each added shell of modes makes the background inferred from the long-distance motion smaller. If the result never stabilizes, there is no cutoff-independent string theory on that background.

9.5 The Metric Must Reach a Fixed Point

Now replace the sphere by a general target-space metric $g_{\mu\nu}(x)$. Integrating in a thin logarithmic shell of oscillator modes changes that metric by a local symmetric rank-two tensor. The change

must vanish in flat space and begin with two derivatives of the metric. At leading order, covariance leaves the Ricci tensor as the relevant candidate:

$$\frac{\partial g_{\mu\nu}}{\partial \log \Lambda} = -C R_{\mu\nu} + \dots, \quad C > 0 \quad (9.11)$$

in the direction of coarse-graining used in the geometric argument. Positive curvature then drives a sphere inward, as required.²

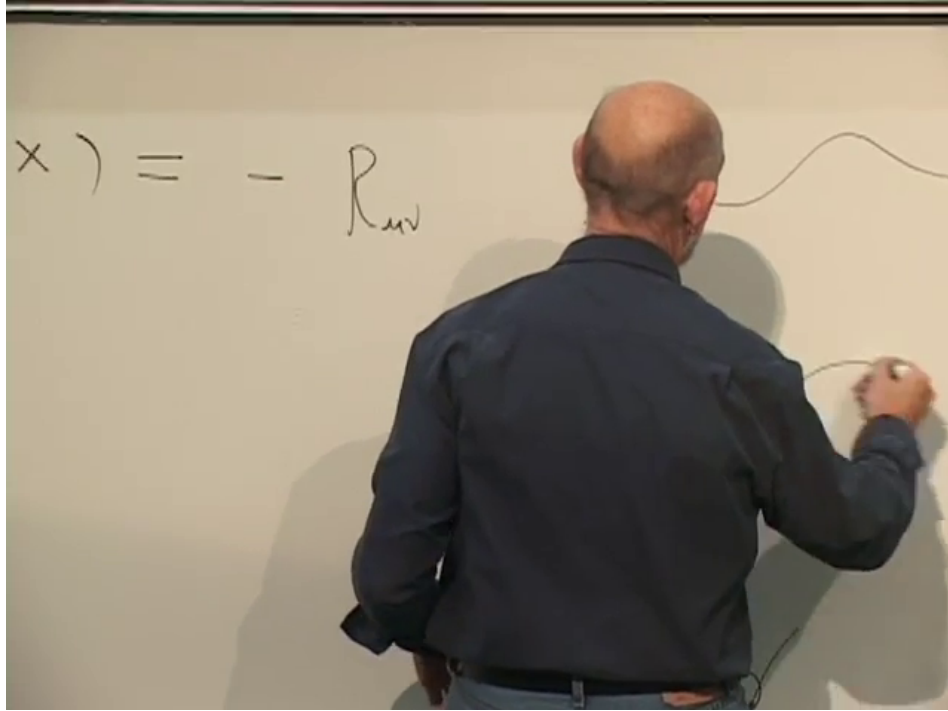


Figure 9.7: The metric change is written schematically as minus the Ricci tensor, with a smoothing curve drawn beside it.

Lecture frame: 00:43:00.

Question from the audience. Where does the oscillator cutoff N appear in this flow equation?

^a

Response. It is the scale variable. The string-size sum changes logarithmically when the highest included mode changes, so the natural evolution parameter is $\log N$, or equivalently $\log \Lambda$. Equation (9.11) describes the infinitesimal change caused by one logarithmic shell of modes.

^aLecture timestamp: 00:43:38.

An admissible background is a fixed point of this flow. To leading order,

$$\boxed{R_{\mu\nu} = 0.} \quad (9.12)$$

²Editorial clarification: The normalization and sign depend on whether the renormalization parameter increases toward the ultraviolet or toward the infrared. In conventional worldsheet notation the leading metric beta function is $\beta_{\mu\nu}^g = \alpha' R_{\mu\nu} + \dots$. The invariant statement used here is that cutoff independence requires the beta function, and hence $R_{\mu\nu}$ at this order, to vanish.

Ricci flat does not mean that every component of the Riemann tensor vanishes. Gravitational waves, for example, can have nonzero curvature in vacuum while satisfying $R_{\mu\nu} = 0$.

9.5.1 The Same Condition Is the Vacuum Einstein Equation

With conventional constants suppressed, Einstein's equation is

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = T_{\mu\nu}. \quad (9.13)$$

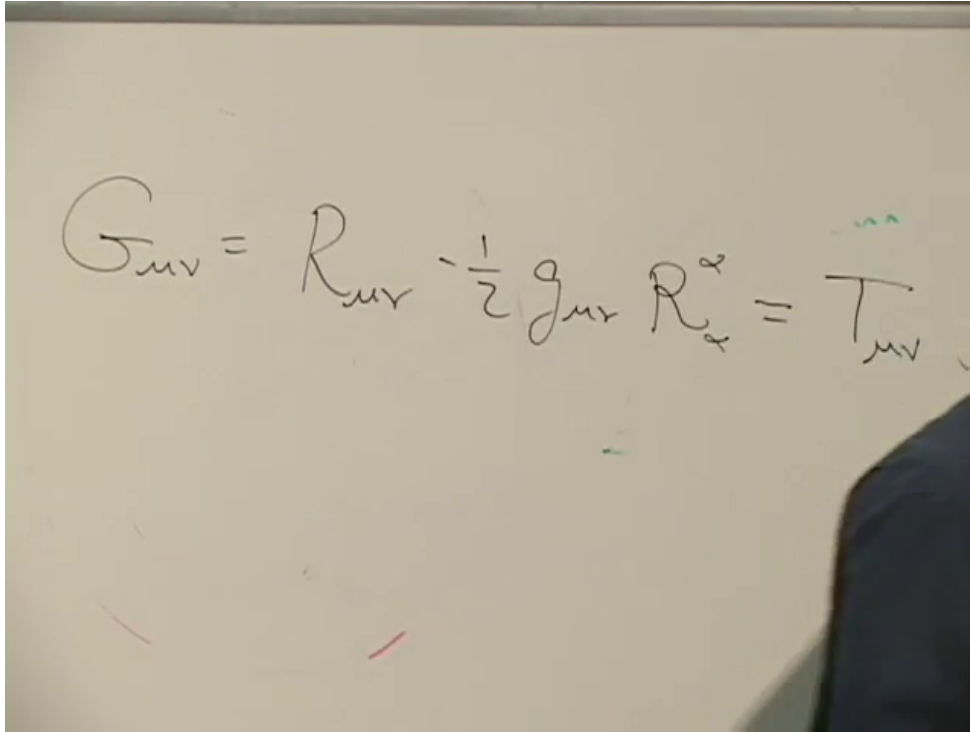


Figure 9.8: The Einstein tensor and stress tensor on the board.

Lecture frame: 00:47:15.

In vacuum $T_{\mu\nu} = 0$. Taking the trace in D spacetime dimensions gives

$$0 = g^{\mu\nu}G_{\mu\nu} = \left(1 - \frac{D}{2}\right)R. \quad (9.14)$$

For $D \neq 2$, $R = 0$, and substitution back into (9.13) yields $R_{\mu\nu} = 0$. Thus the condition that the string lose its cutoff dependence is the vacuum gravitational field equation. Gravity has not been appended to the theory; it has appeared as the price of consistency.

The word *vacuum* does not mean that nothing happens. It means there is no stress-energy source in the region. Source-free Maxwell equations admit electromagnetic waves, and source-free Einstein equations admit gravitational waves. The geometry may be dynamical even when its Ricci tensor vanishes.

Question from the audience. Is this condition connected with conformal invariance? ^a

Response. Yes, although the question is partly lost in the recording. The relevant symmetry is conformal invariance of the *worldsheet* quantum theory. Requiring the worldsheet scale anomaly to vanish is what sets the metric beta function to zero. It should not be confused with demanding conformal invariance of spacetime itself.

^aLecture timestamp: 00:52:20.

There is a second consistency condition. Worldsheet conformal invariance also fixes the critical spacetime dimension: ten for superstrings and twenty-six for the bosonic string. Ricci flatness alone is therefore not the complete list of requirements, but it is the part that displays the Einstein equation most directly.

9.6 Compact Dimensions Rather Than Missing Dimensions

A superstring background has nine spatial dimensions and one time dimension. Our macroscopic world displays only three large spatial directions, so six spatial dimensions must be hidden or otherwise rendered inaccessible at ordinary scales. Compactification keeps them as genuine directions while giving them small finite extent.

Question from the audience. If different observers carry different clocks, does that amount to having more than one time dimension? ^a

Response. No. Many clocks are many instruments measuring coordinates in one spacetime; they do not create independent timelike directions. The extra dimensions considered here are spatial. The familiar quantum and causal constructions do not supply a comparable consistent theory with several local time directions, so that possibility is not part of this compactification model.

^aLecture timestamp: 00:56:59.

Imagine creatures who believe their world is a line. At low resolution they can move only forward and backward. Under a sufficiently powerful microscope, however, the line is revealed to be a very thin cylinder. A small creature can move around its circular direction as well as along its length. The second direction was always present; its circumference was too small for coarse probes to resolve.

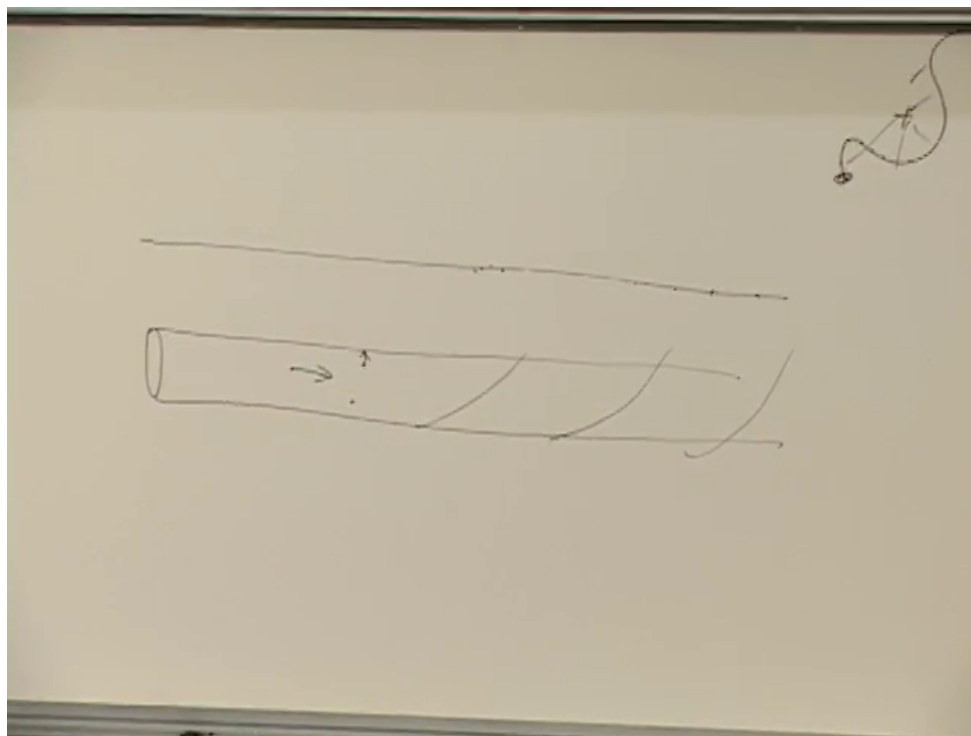


Figure 9.9: A narrow cylinder and trajectories that wind around its hidden circular direction.

Lecture frame: 01:01:00.

A cylinder can be represented as a strip whose upper and lower edges are identified. Crossing the upper edge returns through the lower edge at the same longitudinal position. There is no physical boundary at either edge. The same operation can be performed with two directions: identify the left and right edges of a rectangle and also its top and bottom edges. The result is a two-torus.

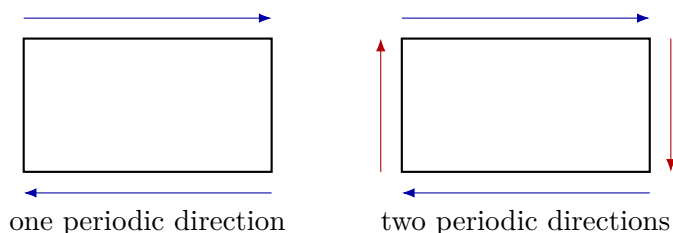
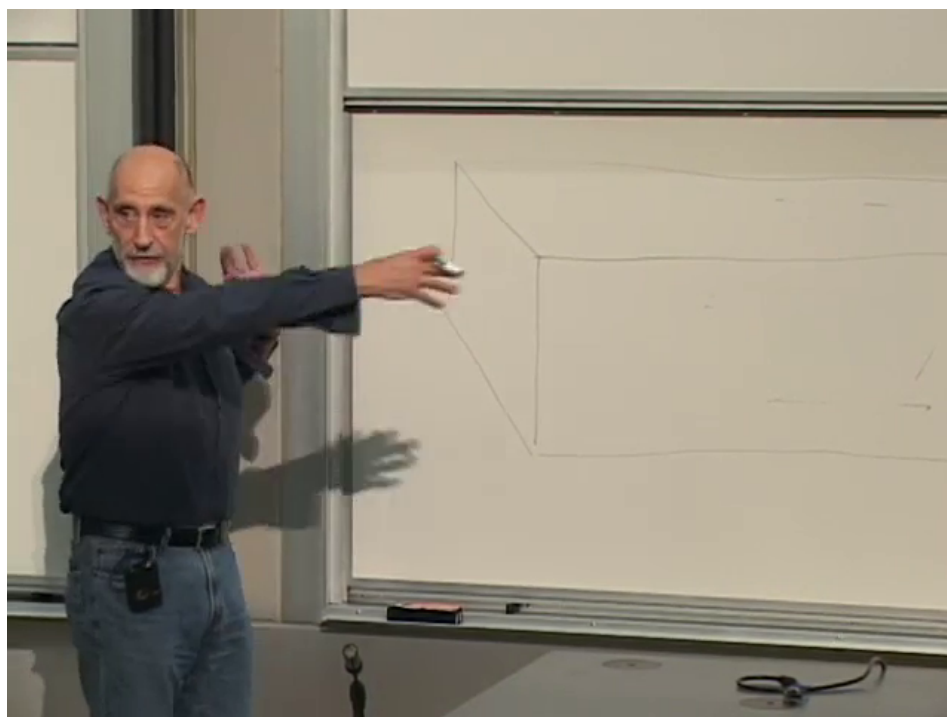


Figure 9.10: The board's slab and a clean fundamental-domain description. Identifying one edge pair gives a cylinder; identifying both pairs gives a torus.

Lecture frame: 01:04:40.

The torus is a topology before it is a bagel-shaped surface in three-dimensional space. Cut a torus once and it becomes a cylinder; cut again and it becomes a rectangle, provided that the edge identifications are remembered. A flat metric may be put on this rectangle, so the resulting flat torus has zero Riemann curvature locally and is certainly Ricci flat.

A straight trajectory on the rectangle continues straight across an identified edge. If its slope is a rational ratio of the two periods, the trajectory eventually closes. If the ratio is irrational, it never closes and becomes dense on the torus.

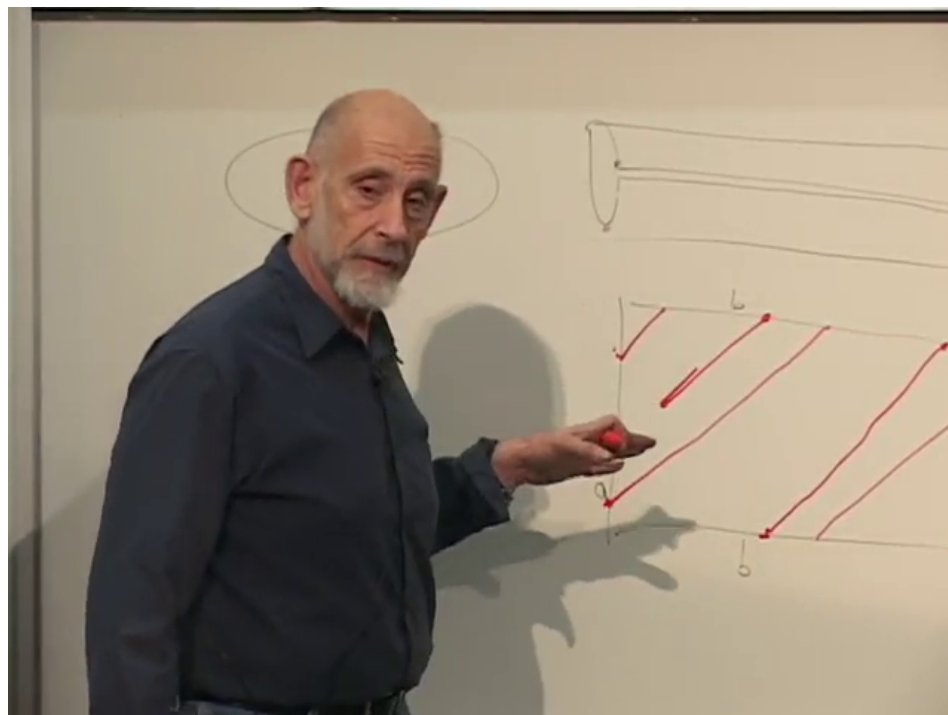


Figure 9.11: Repeated straight segments in a fundamental domain represent one uninterrupted trajectory on the torus.

Lecture frame: 01:11:00.

Higher-dimensional tori are formed in the same way. A cube with each pair of opposite faces identified is a three-torus, and six compact directions can be modeled by a six-torus. One should not try to embed all of these objects in ordinary three-dimensional space; the face-identification description is intrinsic and complete.

Question from the audience. Must an extra dimension necessarily come back on itself? ^a

Response. Not as a theorem. It could be noncompact, but then it would normally appear as another large direction in the observed long-distance world. Compactness is the simplest way to reconcile extra dimensions with their apparent absence. Why nature chooses a particular compactification is a separate dynamical question.

^aLecture timestamp: 01:15:03.

Question from the audience. Why assume the unseen dimensions are ordinary spatial directions curled up in this way rather than something altogether different? ^a

Response. Because this is a definite solvable hypothesis, not because it has been proved to be the unique description of nature. Compact spatial manifolds let us construct consistent examples and compute their spectra. More exotic possibilities must earn their place by doing the same.

^aLecture timestamp: 01:15:51.

Question from the audience. Distances on the familiar bagel-shaped torus do not look like distances in a rectangle. Which geometry is meant? ^a

Response. The term *torus* fixes the identifications, not one unique metric. The embedded bagel inherits a curved metric from three-dimensional space. The periodically identified rectangle carries a flat metric. They have the same topology and different geometry; for the simple compactification we mean the flat one.

^aLecture timestamp: 01:18:11.

9.7 The Shape Parameters of a Flat Torus

Even a flat two-torus is not unique. Its fundamental parallelogram has an overall area, an aspect ratio, and a shear. Changing these parameters changes distances and therefore changes the spectrum, without changing the topology.

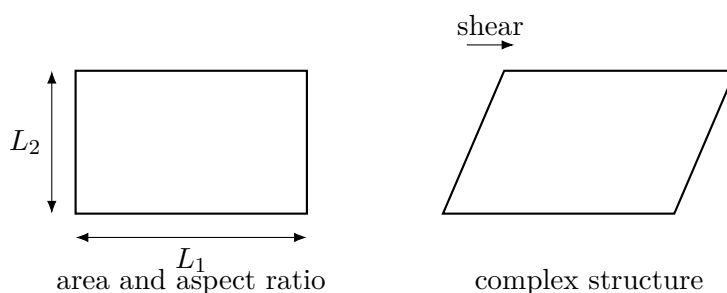
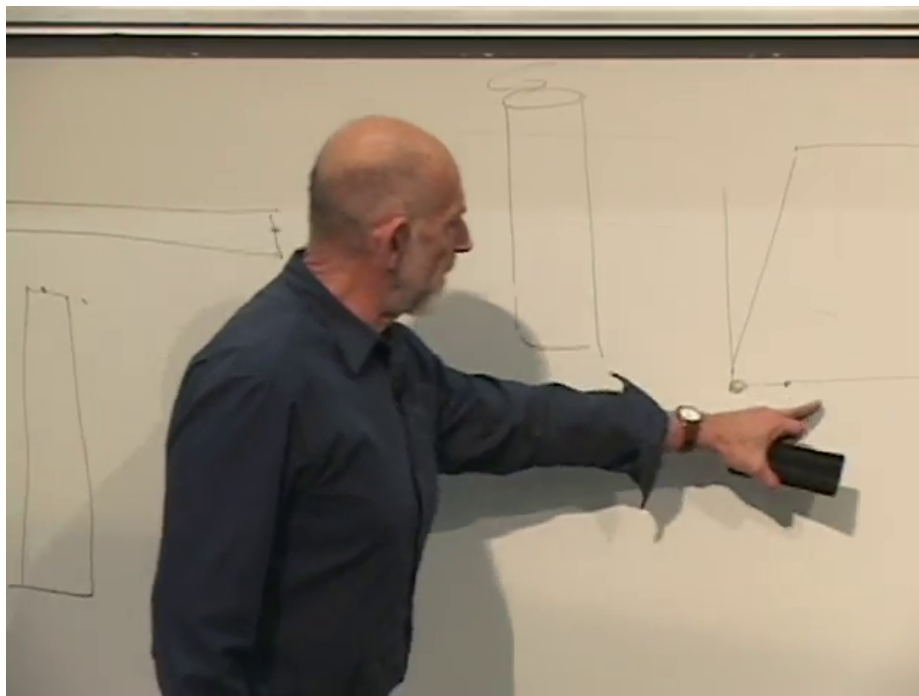


Figure 9.12: The rectangular and sheared fundamental domains used to discuss torus moduli.

Lecture frame: 01:23:00.

For a two-torus it is useful to package the shape into a complex number $\tau = \tau_1 + i\tau_2$: the real part records shear and the imaginary part records an aspect ratio after an overall scale is separated. Interchanging the two basic cycles or changing their integer basis can make apparently different parallelograms describe the same torus.

Question from the audience. What happens if one edge is flipped before the identification?^a

Response. That produces a different topology. Identifying with a reflection gives a nonorientable surface, the Klein bottle, rather than an ordinary torus.

^aLecture timestamp: 01:24:20.

Question from the audience. What if the twist is not uniform along the edge?^a

Response. A smooth nonuniform deformation can be transferred into the metric on the fundamental domain. The topological identification and the geometrical data must be kept distinct: changing the gluing may change topology, while smoothly changing distances changes the metric.

^aLecture timestamp: 01:25:01.

3

Flat tori are useful because they solve the consistency condition immediately, but their symmetry is too generous for a realistic model of particle physics. More intricate Ricci-flat spaces, notably Calabi–Yau manifolds, provide richer compactifications. Their detailed construction lies beyond this discussion; the principle remains the same.

9.8 Reading a Compact Circle from Its Spectrum

After the break, reduce the compact space to one circle of circumference $2\pi R$. A wavefunction must be single-valued after one circuit:

$$e^{ip_c(x+2\pi R)/\hbar} = e^{ip_c x/\hbar}. \quad (9.15)$$

Hence

$$p_c R = n\hbar, \quad p_c = \frac{n\hbar}{R}, \quad n \in \mathbb{Z}. \quad (9.16)$$

³Editorial clarification: For a flat two-torus, the area is the simplest Kahler modulus and τ is the complex-structure modulus. A continuous shear is part of τ . A Dehn twist is instead a discrete mapping-class operation, represented by $\tau \mapsto \tau + 1$; it is not itself the continuous modulus. This sharpens the intentionally informal terminology used during the board discussion.

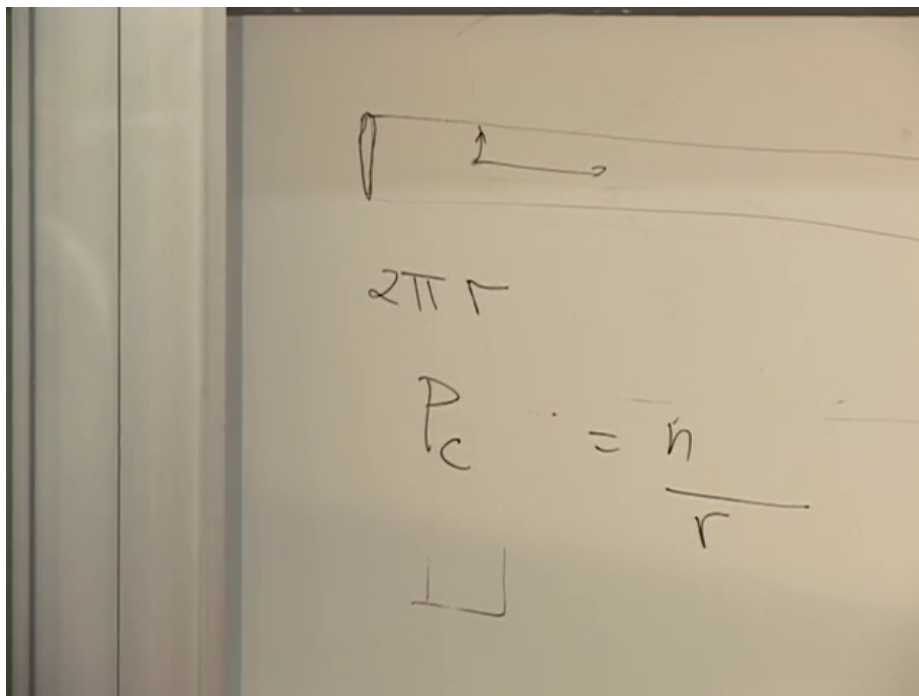


Figure 9.13: The compact circumference $2\pi R$ and quantized momentum $p_c = n\hbar/R$.

Lecture frame: 01:32:15.

For a massless particle in the full spacetime,

$$E = c|\mathbf{P}|. \quad (9.17)$$

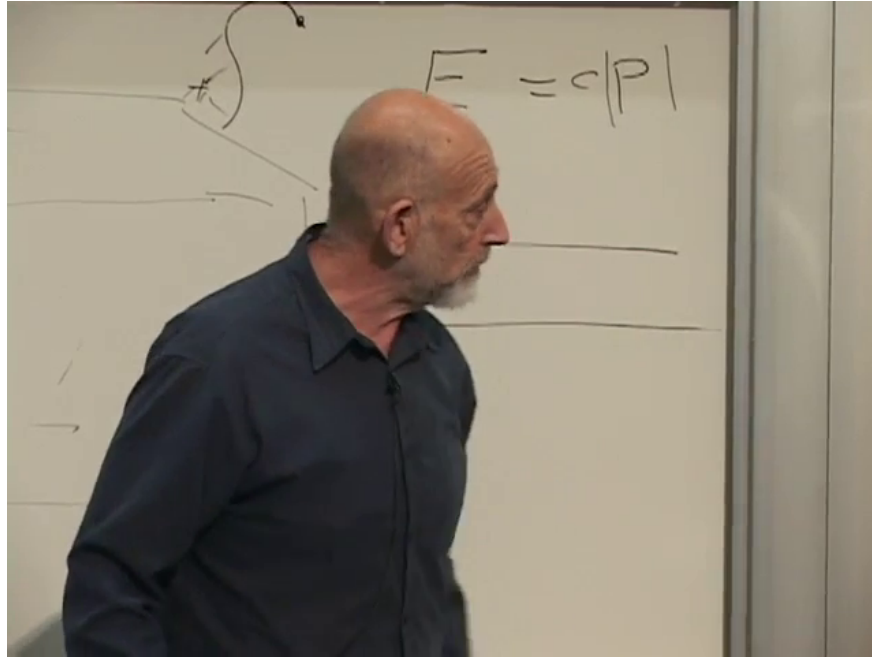


Figure 9.14: The massless dispersion relation used to reinterpret compact momentum as lower-dimensional mass.

Lecture frame: 01:33:45.

Split the momentum into noncompact and compact parts and set $c = \hbar = 1$:

$$E^2 = \mathbf{p}_{\text{large}}^2 + \frac{n^2}{R^2}. \quad (9.18)$$

An observer who cannot resolve the compact direction compares this with $E^2 = \mathbf{p}_{\text{large}}^2 + m^2$ and infers a tower of masses

$$m_n = \frac{|n|}{R}. \quad (9.19)$$

The two signs of n describe opposite compact momenta with the same positive energy, not positive and negative masses. The spacing $1/R$ therefore reveals the radius spectroscopically. This is the Kaluza–Klein mechanism: momentum in a hidden direction appears as mass in the visible dimensions.

9.9 Winding States Complete the Spectrum

A closed string has a possibility unavailable to a point particle: it can wrap around the circle. If the winding number is w , the string closes only after covering the compact direction w times. Its length is $2\pi|w|R$, so tension supplies an energy

$$m_w = 2\pi T |w|R = \frac{|w|R}{\alpha'}. \quad (9.20)$$

The lecture chooses units with $\alpha' = 1$, reducing this to

$$m_w = |w|R. \quad (9.21)$$

Positive and negative w are opposite orientations around the circle.

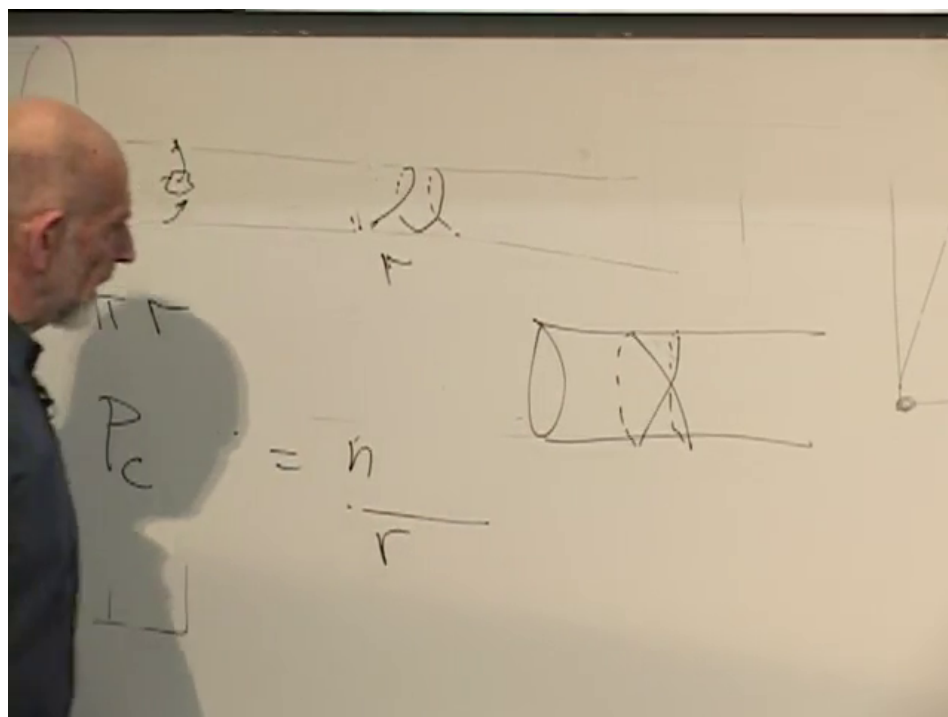


Figure 9.15: A closed string wound around the compact cylinder, with the compact radius marked on the board.

Lecture frame: 01:39:45.

The compact circle thus produces two complementary ladders:

$$m_n = \frac{|n|}{R}, \quad m_w = \frac{|w|R}{\alpha'}. \quad (9.22)$$

When R is large, momentum levels are closely spaced and winding states are heavy. When R is small, momentum levels are widely spaced and winding states are light.

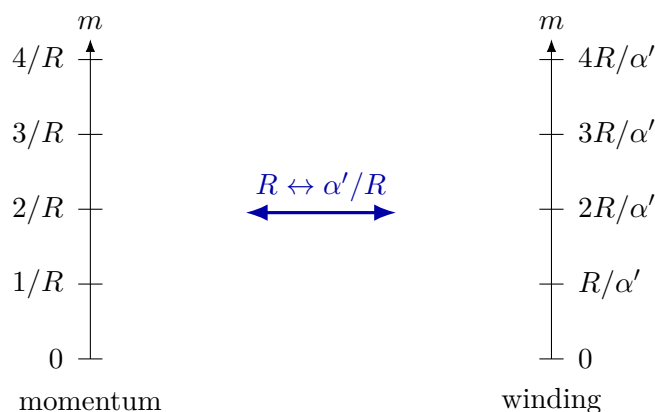


Figure 9.16: Momentum and winding towers exchange when the radius is inverted in string units.

9.10 T-Duality: Small and Large Circles Are the Same Physics

The two towers are interchanged by

$$R \longleftrightarrow \frac{\alpha'}{R}, \quad n \longleftrightarrow w. \quad (9.23)$$

In the units used on the board, $\alpha' = 1$, so this is simply $R \leftrightarrow 1/R$. At the self-dual radius $R = \sqrt{\alpha'}$, momentum and winding scales coincide.

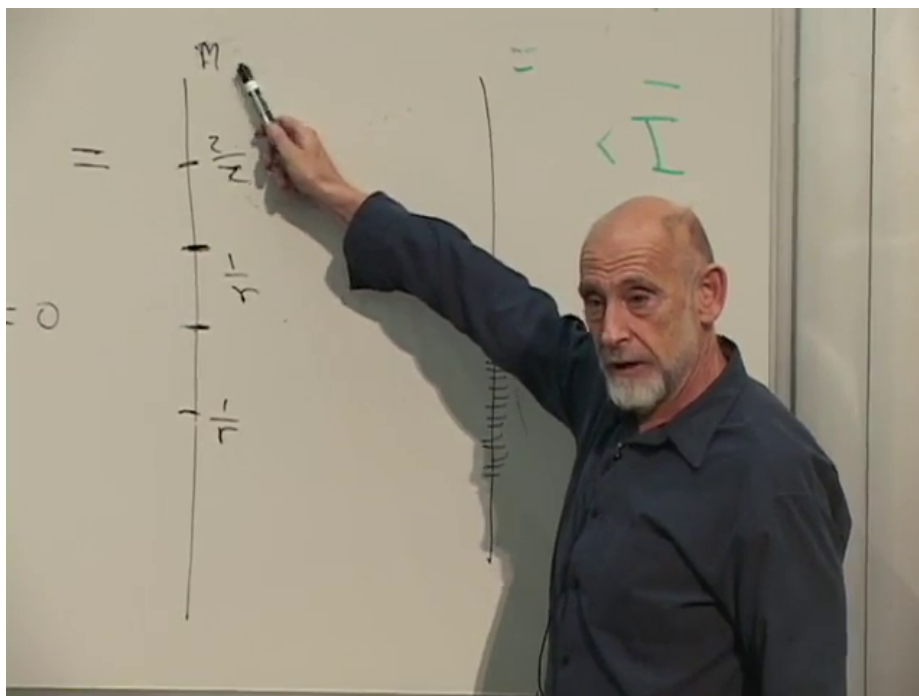


Figure 9.17: The sparse and dense mass ladders compared on the board as the radius crosses the self-dual scale.

Lecture frame: 01:43:00.

The exchange of the elementary mass formulas is only the visible edge of the result. Closed-string oscillator states and scattering amplitudes are also mapped consistently. A circle smaller than the string scale is not a new increasingly tiny geometry; it has an equivalent description as a larger dual circle. Point particles cannot make this identification because they possess momentum modes but no winding modes.

Question from the audience. Is this merely the same as exchanging a vertical and a horizontal direction? ^a

Response. No. It exchanges two different quantum numbers: momentum around the circle and the number of times a closed string wraps the circle. The radius must be inverted at the same time.

^aLecture timestamp: 01:44:38.

Question from the audience. Do the descriptions at $R < \sqrt{\alpha'}$ and $R > \sqrt{\alpha'}$ refer only to bosons, or also to fermions? ^a

Response. The duality is a symmetry of the full compactified string theory, not just the simple bosonic mass ladder used to expose it. The detailed mapping of sectors depends on the theory, but the small-large equivalence is not restricted to one particle species.

^aLecture timestamp: 01:46:30.

Question from the audience. Are the two states being compared open strings or closed strings? ^a

Response. They are closed-string states. A freely propagating closed string can wind a compact circle. An ordinary open string with free endpoints does not carry the same conserved winding number.

^aLecture timestamp: 01:47:32.

4

Question from the audience. Is the interchange analogous to the exchange of space and time near a black-hole horizon? ^a

Response. No. Horizon coordinates and causal character are a different issue. Here the two descriptions concern one compact spatial circle and the exchange of its momentum and winding sectors.

^aLecture timestamp: 01:48:00.

Question from the audience. Does this cure the short-distance self-energy problem by giving a minimum length? ^a

Response. It shows that radii below the string scale do not define new inequivalent circle backgrounds: decreasing R beyond the self-dual point increases the dual radius. This is a precise stringy obstruction to probing arbitrarily small circle geometry, although it should not be turned into a claim that every conceivable notion of distance has one universal hard lower bound.

^aLecture timestamp: 01:48:35.

9.11 Why Winding Sectors Cannot Be Omitted

Question from the audience. Could we define the theory using only unwound strings and leave winding states out? ^a

Response. No. Closed-string interactions generate them. Stretch an unwound loop around the compact direction so that one segment follows the circle one way and another follows it the opposite way. The net winding is still zero. A reconnection can then split the loop into two closed strings with winding $+1$ and -1 . The pair may separate, so winding states must belong to the spectrum even if the initial state had none.

^aLecture timestamp: 01:48:56.

⁴Editorial clarification: In the fuller open-string account, T-duality exchanges Neumann and Dirichlet boundary conditions and introduces D-branes. That is why open-string T-duality is richer than simply copying the closed-string winding formula.

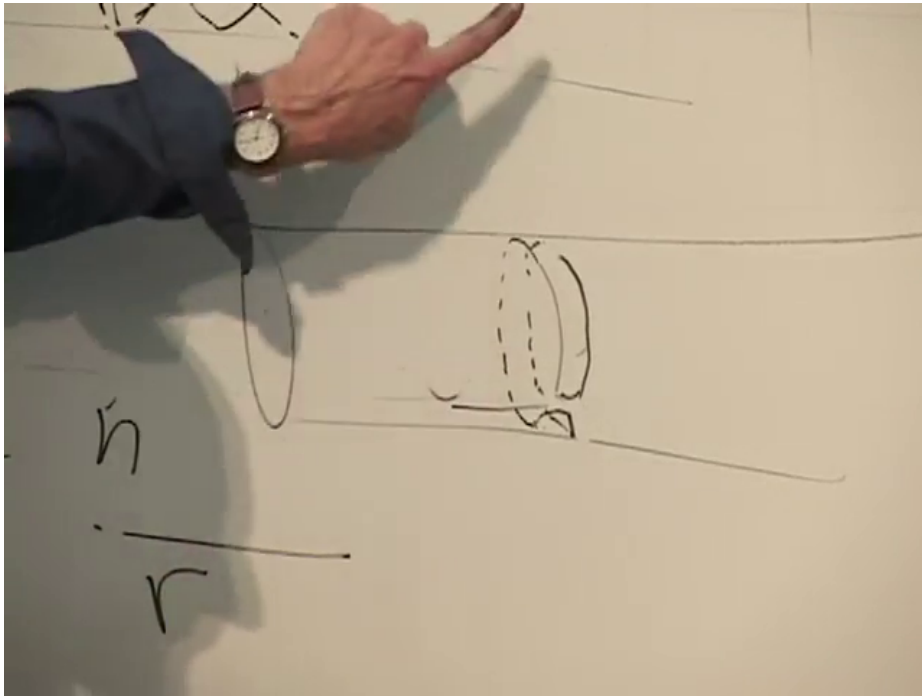


Figure 9.18: The stretched zero-winding loop arranged around the compact direction before reconnection.

Lecture frame: 01:50:30.

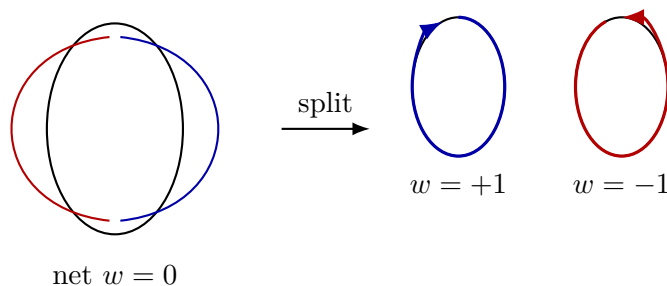
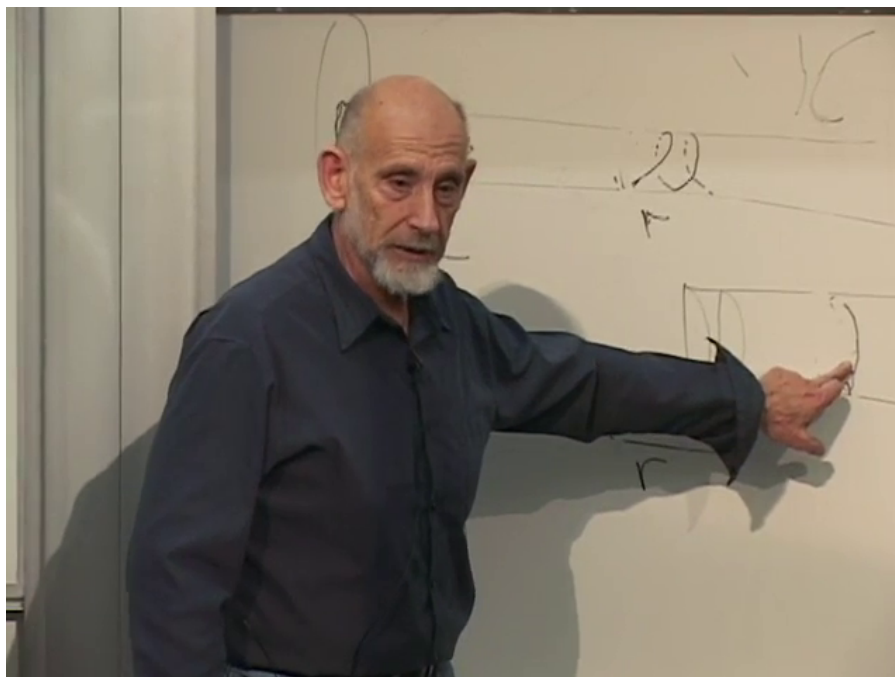


Figure 9.19: The board's separated wound loop and a clean reconstruction of winding–antiwinding pair production. Total winding remains zero.

Lecture frame: 01:51:30.

Question from the audience. Does producing the winding pair require extra energy? ^a

Response. Yes. Stretching a string costs tension energy, and the final wound strings have the masses in (9.20). A sufficiently energetic collision can produce the pair, just as a collision can produce a particle and its antiparticle while preserving the relevant conserved charge.

^aLecture timestamp: 01:52:11.

9.12 Can One Hear the Shape of a String Background?

For an ordinary compactified particle theory, the spacing of the Kaluza–Klein spectrum lets us infer the radius: one tries to hear the shape of the compact space from its normal modes. String theory makes the inverse problem subtler. Momentum modes hear $1/R$, winding modes hear R/α' , and the complete spectrum is unchanged when the two are exchanged. Two radii that point-particle geometry regards as different are one string background described in dual variables.

The lecture has therefore arrived at both of its promised conclusions by one route. Ultraviolet consistency forces the target metric to satisfy the vacuum Einstein equation, so gravity is unavoidable. Compactification then turns hidden geometry into particle spectra, but winding states prevent that spectrum from assigning a unique classical radius below the string scale. Consistency does not merely select which geometries are allowed; through T-duality, it changes what counts as a distinct geometry at all.

Chapter 10

T-Duality, D-Branes, and Gauge Theory

Overview. T-duality begins as an exact equivalence between closed strings on circles of radii (R) and α'/R . Momentum and winding exchange roles, as do the worldsheet derivatives $\partial_\tau Y$ and $\partial_\sigma Y$, and the spacetime vectors descending from $G_{\mu 5}$ and $B_{\mu 5}$. When the same operation is applied to an open string, a Neumann endpoint condition becomes a Dirichlet condition. The endpoint must then lie on a dynamical D-brane. Open strings confined to one brane produce an abelian gauge theory at low energy; strings joining a stack of branes carry two endpoint labels and produce a nonabelian gauge theory. A D-string ending on a D3-brane supplies the corresponding magnetic monopole.

10.1 Why Return to T-Duality?

We have already met T-duality, but we return to it for a reason. It was one of the clues that string theory contained much more than perturbative strings. Once its consequences for open strings were understood, it led to D-branes; D-branes in turn became a way to construct and study quantum field theories. The aim of this lecture is therefore not merely to repeat the slogan that a small circle is equivalent to a large circle. We shall rebuild the argument carefully enough to see how that slogan forces new physical objects into the theory.

The simplest setting is toroidal compactification. A one-dimensional torus is a circle. A two-dimensional torus can be represented by a parallelogram whose opposite sides are identified, and a three-dimensional torus by a parallelepiped whose opposite faces are identified. The same construction continues in higher dimensions. We shall keep only one compact spatial coordinate, called Y , and one representative noncompact coordinate, called X . Time is present but is not shown in the board diagrams.

Choose the standard convention

$$Y \sim Y + 2\pi R, \quad (10.1)$$

so that R is the radius and $2\pi R$ is the circumference. The lecture often calls the compact size simply r and suppresses factors of 2π and α' . We shall retain the conventional factors in the derivations and display the lecture units when they are useful.

For the first half of the discussion the strings are closed and oriented. An orientation is an intrinsic

arrow drawn around the string. Reversing all arrows everywhere would change no experiment; what matters is the relative orientation of two strings. During splitting and joining, the arrows must join continuously. This is the first hint that a signed, conserved quantity will appear.

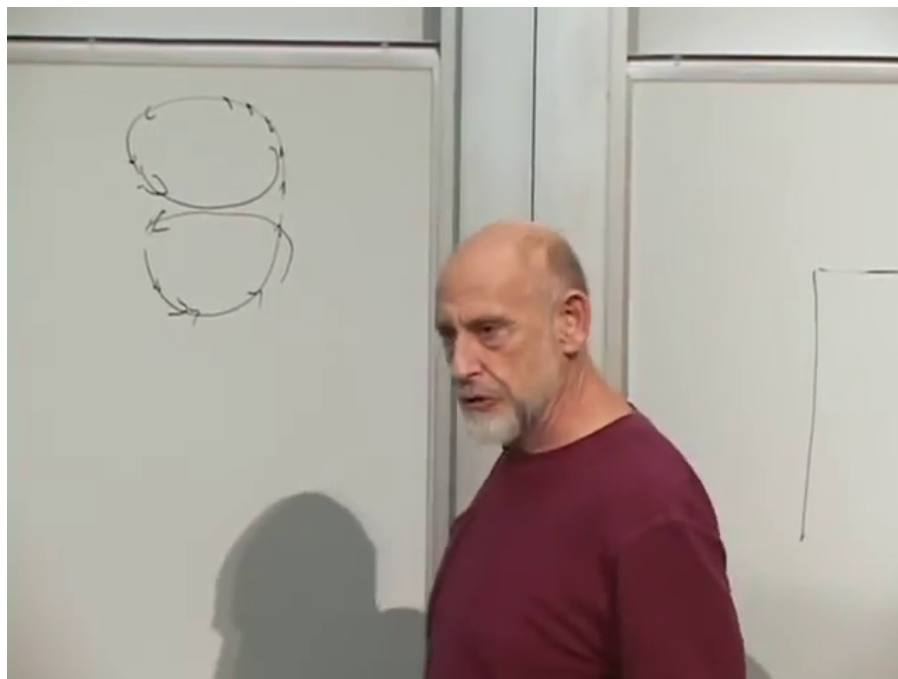


Figure 10.1: Two closed strings with their relative orientations marked on the board. Orientation must be continuous through splitting and joining.

Lecture frame: 00:05:40.

Question from the audience. If the worldsheet coordinate runs from $\sigma = 0$ to 2π , has the string already gone once around the compact direction? ^a

Response. No. The coordinate σ labels points along the string; it does not label positions in the compact space. Every closed string has a periodic parameter σ , whether it is wound or unwound. The rubber-band analogy is useful: one may number the molecules around a rubber band even when the band is not wrapped around anything. Winding is a property of the map $Y(\tau, \sigma)$, not of the parameter σ alone.

^aLecture timestamp: 00:08:42.

Question from the audience. Must the compact size also be the same at every value of the large coordinate X ? ^a

Response. Not in a general dynamical compactification. We begin with a fixed circle because it is the cleanest background. Later the component G_{55} will become a scalar field whose value measures the local size of the fifth direction, so that size can vary over ordinary spacetime.

^aLecture timestamp: 00:10:00.

10.2 Winding Is Conserved Even When Strings Reconnect

An unwound string may move through the compact direction without wrapping it. A wound string is different: as σ goes once around the closed string, the embedding $Y(\tau, \sigma)$ advances through one or more periods of the compact circle. Because the string is oriented, the advance has a sign. A string can wind in the positive direction or in the negative direction.

The conservation law is clearest in the joining pictures. Take one string with winding $+1$ and another with winding -1 . Their compatible segments may reconnect. The resulting configuration has zero net winding and can be deformed into an unwound string. The board traces precisely this opposite-orientation reconnection.

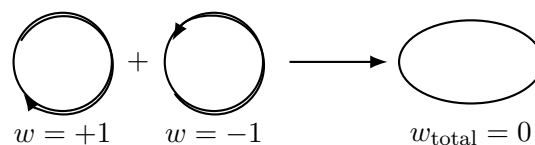
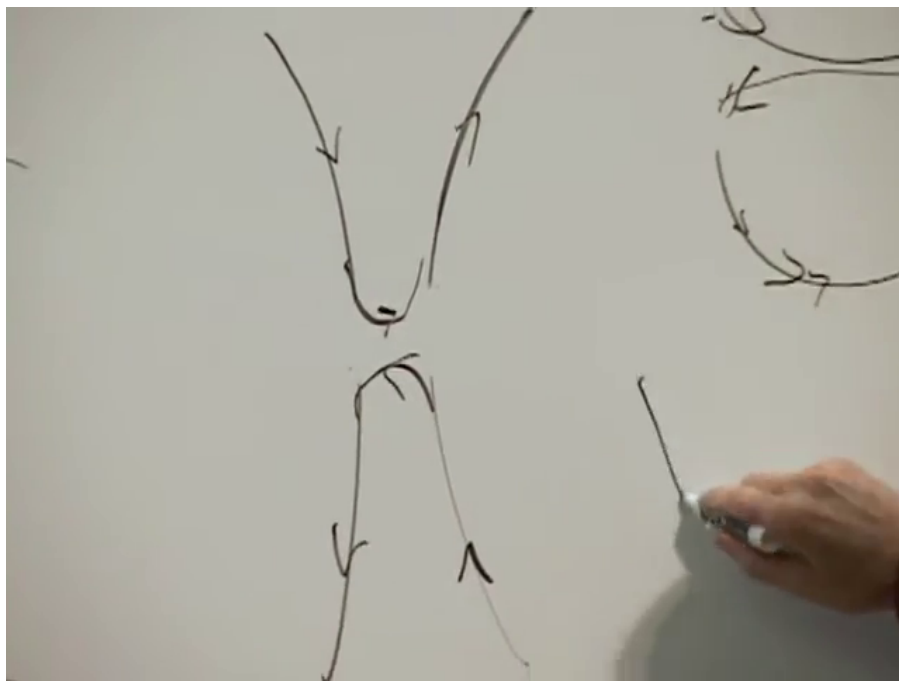


Figure 10.2: Opposite windings reconnect to a zero-net-winding configuration. The lower diagram is a clean reconstruction of the board move, not a second dynamical process.

Lecture frame: 00:13:20.

Now repeat the experiment with two strings having the same orientation. They can join, but they cannot unwind. The result carries twice the winding of either incoming string. The number of connected strings has changed; the total winding has not.

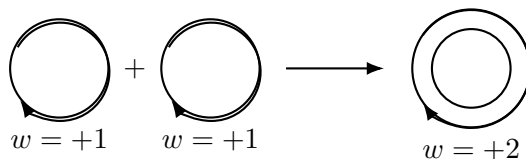
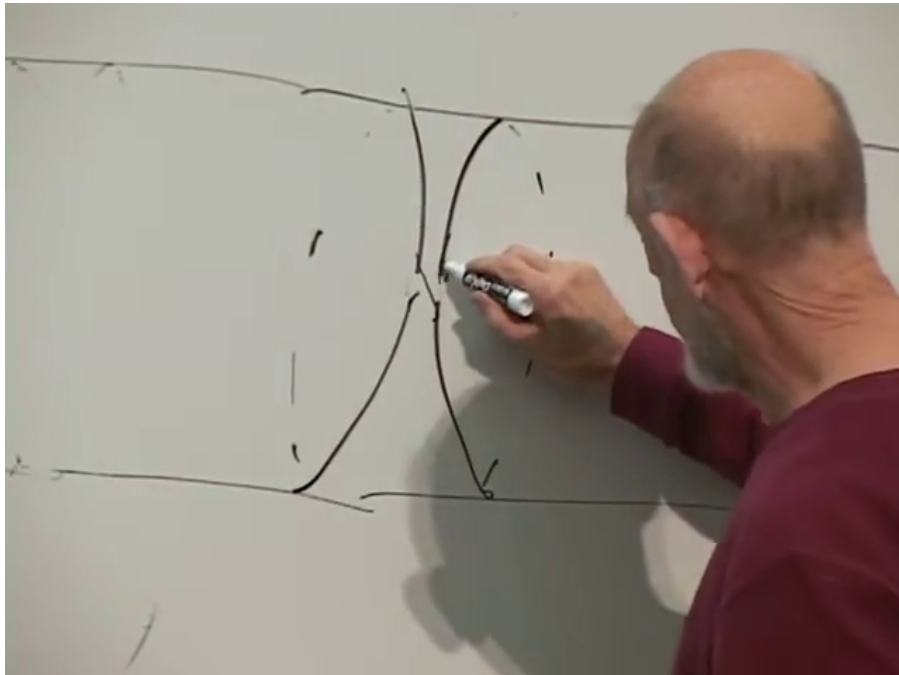


Figure 10.3: Same-sign windings can reconnect into a doubly wound string. Connectivity changes while total winding remains two.

Lecture frame: 00:16:40.

This distinction survives every allowed process. A total winding of one hundred might be carried by one string wound one hundred times, two strings wound fifty times each, or one hundred singly wound strings. Strings can split and join among these descriptions, but the signed sum remains one hundred. We denote it by $w \in \mathbb{Z}$.

10.3 Momentum Modes and Winding Modes

Conservation by itself is not yet duality. We next compare energies. A state moving around the compact circle has quantized momentum

$$p_Y = \frac{n}{R}, \quad n \in \mathbb{Z}. \quad (10.2)$$

For a state whose other mass contributions can be neglected, the compact motion contributes $E_{\text{mom}} = |n|/R$. Positive and negative n denote opposite directions of motion, while the energy is positive.

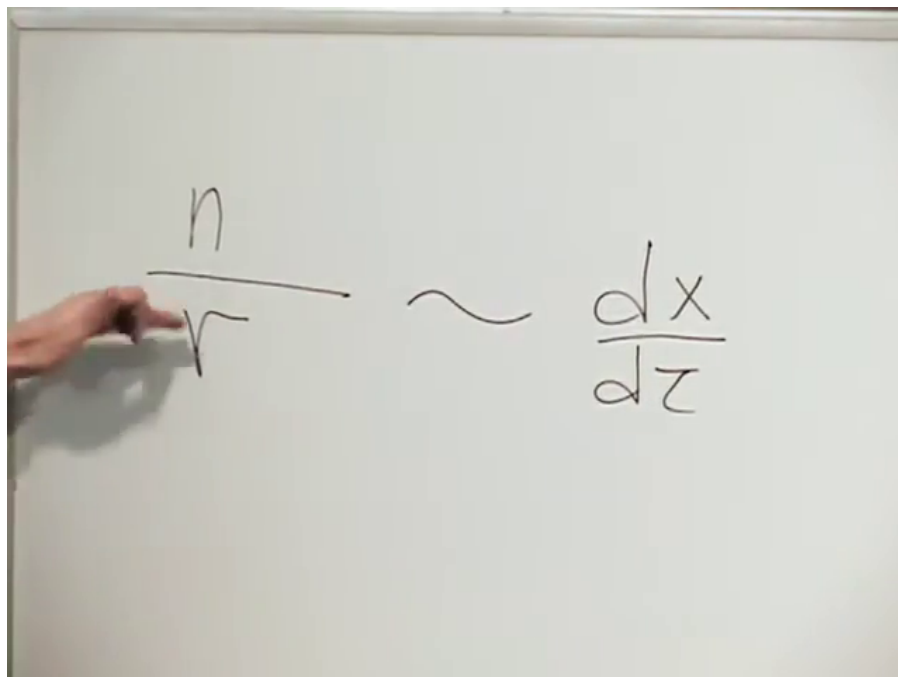


Figure 10.4: The compact momentum quantum number on the board. The lecture writes the momentum schematically as an integral of the compact velocity.

Lecture frame: 00:21:30.

A string wound $|w|$ times has length $2\pi|w|R$. With fundamental string tension

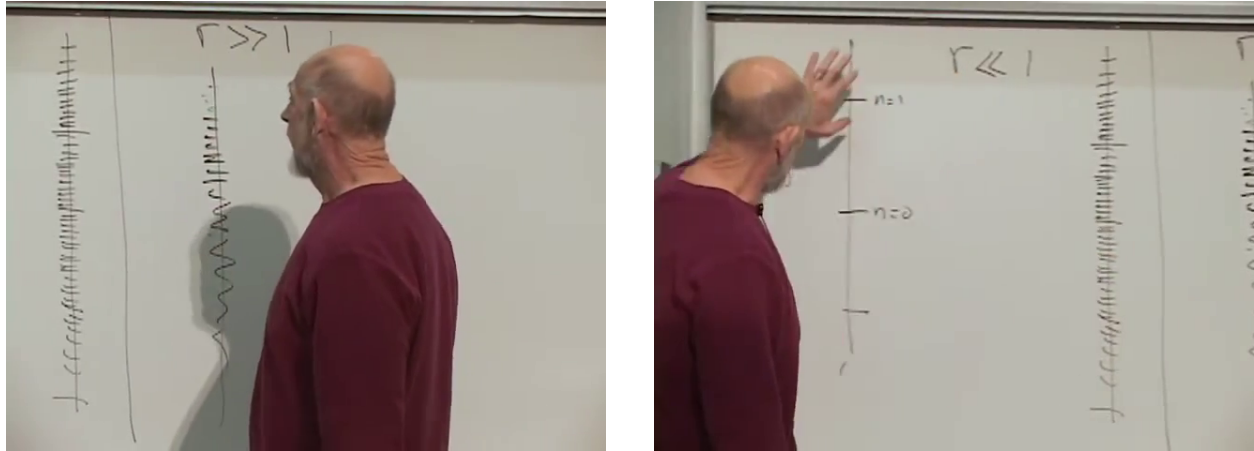
$$T_F = \frac{1}{2\pi\alpha'}, \quad (10.3)$$

its winding energy is

$$E_{\text{wind}} = T_F(2\pi|w|R) = \frac{|w|R}{\alpha'}. \quad (10.4)$$

In the lecture units $\alpha' = 1$, this is simply $E_{\text{wind}} = |w|R$.

The two towers behave oppositely as the radius changes. For $R \gg \sqrt{\alpha'}$, momentum levels are closely spaced and winding levels are widely spaced. For $R \ll \sqrt{\alpha'}$, momentum levels are widely spaced and winding levels are closely spaced.



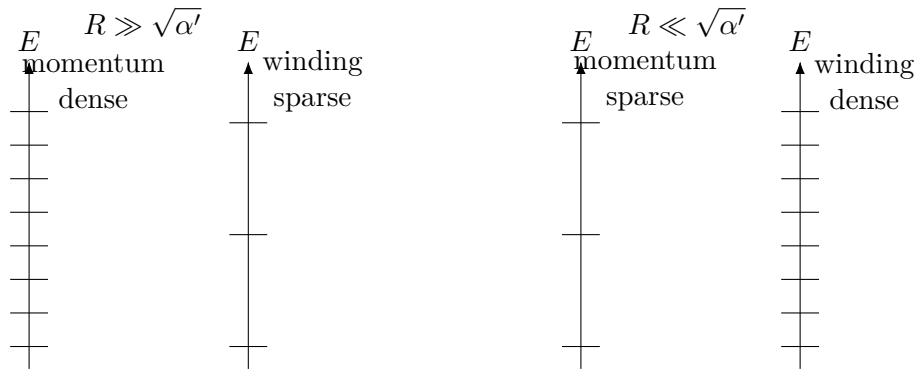
Large compact radius.

Small compact radius.

Figure 10.5: The two spectral comparisons on the board. Sparse and dense towers exchange roles as the radius is inverted.

Lecture frame: 00:25:30–00:26:00.

The clean schematic is


 Figure 10.6: Reconstruction of the level-spacing argument. The complete string spectrum is unchanged when momentum and winding are exchanged together with $R \leftrightarrow \alpha'/R$.

Suppose an observer is handed one of these spectra without being told how it was made. Is the closely spaced tower momentum on a large circle, or winding on a small circle? The spectrum alone cannot say. String theory makes the stronger statement that scattering amplitudes cannot say either. The two descriptions are exactly equivalent under

$$n \longleftrightarrow w, \quad R \longleftrightarrow \tilde{R} = \frac{\alpha'}{R}. \quad (10.5)$$

The fixed point is

$$R = \sqrt{\alpha'}, \quad (10.6)$$

or $R = 1$ in the lecture units. At that radius additional states can become massless and the gauge symmetry is enhanced.

Question from the audience. Is winding only a mathematical convenience for rewriting the spectrum? ^a

Response. No. Winding labels real sectors of the closed-string Hilbert space, and those sectors take part in interactions. The word “duality” means that the momentum description and winding description are different coordinates on the same physics, not that one sector is an optional mnemonic for the other.

^aLecture timestamp: 00:27:33.

Question from the audience. Is there a radius that is mapped to itself? ^a

Response. Yes. The self-dual value is $R = \sqrt{\alpha'}$, which is $R = 1$ in the units used on the board. The two towers meet there, and the spectrum has enhanced symmetry.

^aLecture timestamp: 00:27:58.

10.4 The Worldsheet Form of the Duality

The spectral argument is suggestive, but the worldsheet makes the exchange more precise. In conformal gauge, the momentum conjugate to Y is proportional to $\partial_\tau Y$. Thus the total compact momentum has the form

$$P_Y = \frac{1}{2\pi\alpha'} \int_0^{2\pi} d\sigma \partial_\tau Y. \quad (10.7)$$

The winding number is

$$w = \frac{1}{2\pi R} \int_0^{2\pi} d\sigma \partial_\sigma Y = \frac{Y(2\pi) - Y(0)}{2\pi R}. \quad (10.8)$$

Question from the audience. Why does the integral of a derivative not vanish around the closed string? ^a

Response. Because Y is a periodic target-space coordinate, not a single-valued real function on the σ -circle. The worldsheet point $\sigma = 2\pi$ is the same as $\sigma = 0$, but a lift of the embedding to the covering line changes by $2\pi w R$. The integral of $\partial_\sigma Y$ measures that shift.

^aLecture timestamp: 00:32:33.

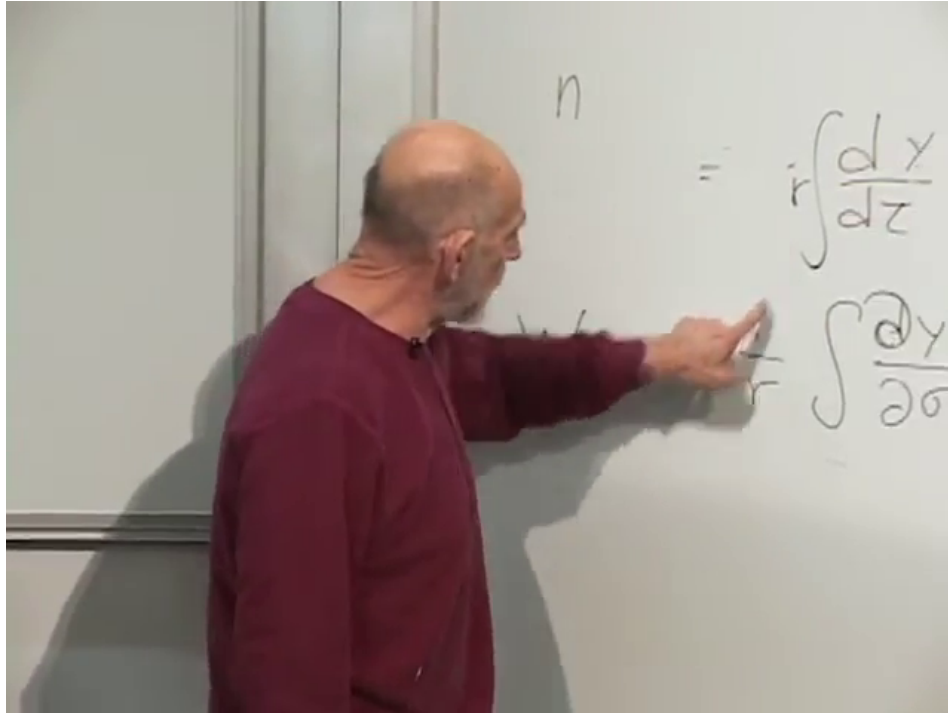


Figure 10.7: The momentum and winding integrals after both lines have been completed on the board. Common 2π and α' factors are suppressed in the lecture normalization.

Lecture frame: 00:33:25.

For a once-wound string,

$$Y(\tau, \sigma + 2\pi) = Y(\tau, \sigma) + 2\pi R. \quad (10.9)$$

For winding w , the right-hand shift is $2\pi w R$. Equation (10.8) then counts the winding exactly. This is the precise version of the board shorthand $w \sim R^{-1} \int d\sigma \partial_\sigma Y$.

The worldsheet statement of T-duality is therefore

$$\partial_\tau Y \longleftrightarrow \partial_\sigma \tilde{Y}, \quad \partial_\sigma Y \longleftrightarrow \partial_\tau \tilde{Y}, \quad (10.10)$$

with momentum in Y becoming winding in the dual coordinate $\tilde{Y}$, and conversely.

Question from the audience. Does this mean that the worldsheet coordinates τ and σ themselves are related or interchanged? ^a

Response. No. T-duality maps solutions and exchanges particular derivatives of the compact target coordinate. It does not identify the worldsheet time coordinate with the worldsheet spatial coordinate.

^aLecture timestamp: 00:34:39.

10.5 Why Winding States Cannot Be Avoided

One might ask why a string would ever be found in a wound state when an unwound state is available. A string does not choose its state by preference. Energy supplied in a collision stretches it.

A sufficiently energetic, highly distorted string can reconnect into two strings with opposite winding. Their total winding remains zero, but after the split the two strings may separate. One member of the pair can be sent far away, leaving an observer with an apparently isolated wound string.

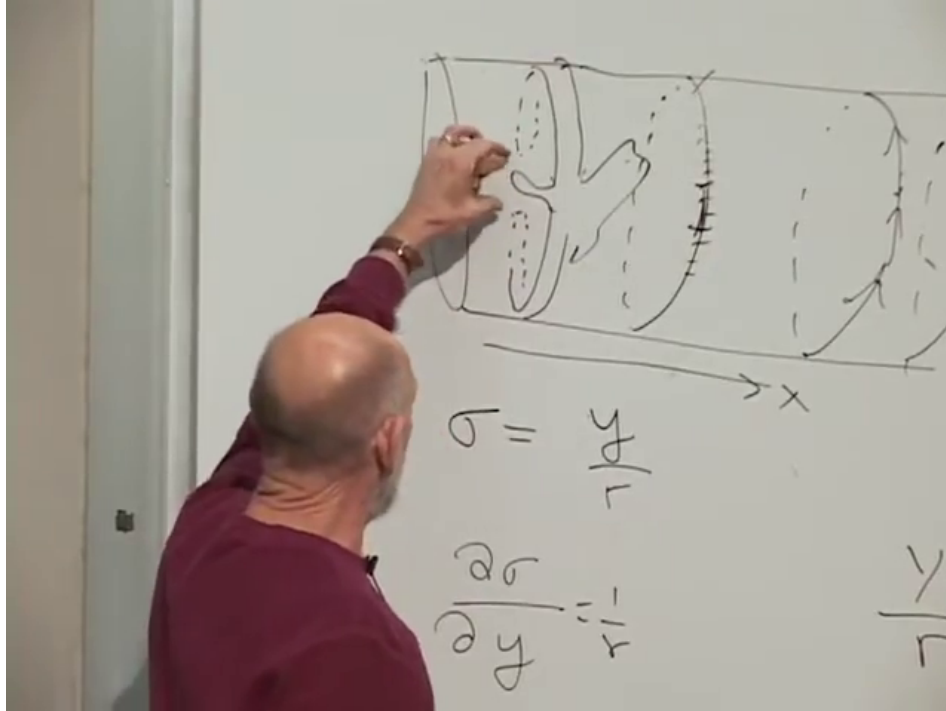


Figure 10.8: A highly stretched zero-net-winding configuration from which oppositely wound strings can separate after reconnection.

Lecture frame: 00:36:15.

The energy threshold is set by the winding energy. On a large circle it takes a large collision energy to stretch enough string around the compact direction. Above threshold, however, winding–antiwinding pair production cannot be forbidden. The analogy is photon collisions producing an electron–positron pair. The total electric charge is still zero, yet the charged particles are present and can be separated. Winding conservation works in the same way.

This analogy is more than verbal. Compact momentum and winding behave as two distinct charge quantum numbers. Like signs repel and opposite signs attract in the corresponding reduced description. We should therefore ask for the two vector fields that mediate those forces.

10.6 Momentum Charge from Five-Dimensional Gravity

Begin with gravity in five dimensions. Let capital indices M, N run over the four ordinary spacetime directions and the compact fifth direction. The metric decomposes as

$$G_{MN} \longrightarrow \{G_{\mu\nu}, G_{\mu 5}, G_{55}\}. \quad (10.11)$$

The symmetric component $G_{5\mu}$ contains no new information beyond $G_{\mu 5}$.

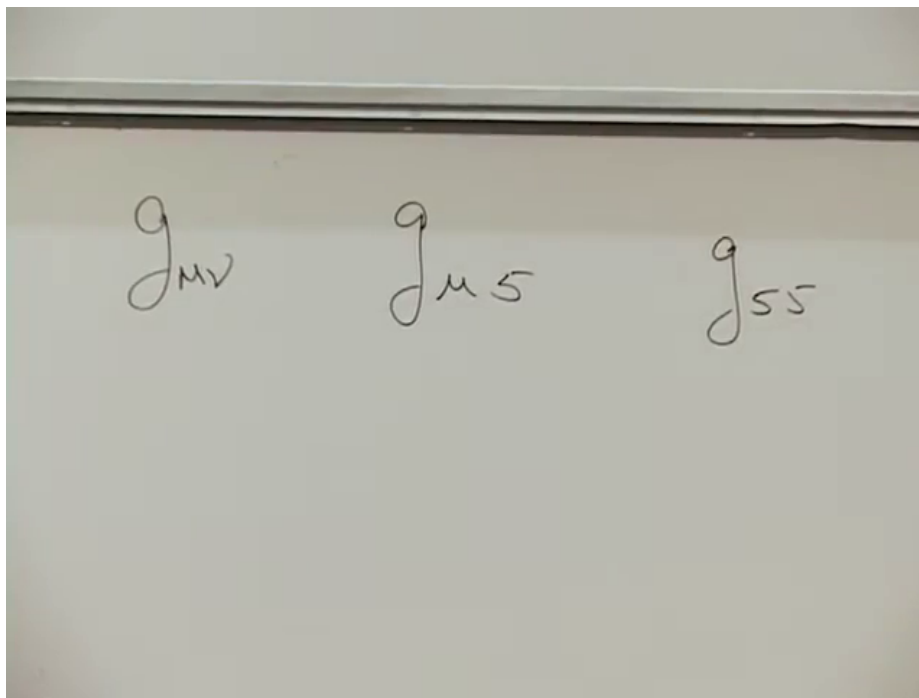


Figure 10.9: The independent groups of five-dimensional metric components: $G_{\mu\nu}$, $G_{\mu 5}$, and G_{55} .

Lecture frame: 00:44:05.

From the four-dimensional point of view,

$$G_{\mu\nu} \text{ is a spacetime metric,} \quad (10.12)$$

$$G_{\mu 5} \equiv A_\mu \text{ is a vector field,} \quad (10.13)$$

$$G_{55} \equiv e^{2\varphi} \text{ is a scalar modulus.} \quad (10.14)$$

The scalar φ measures the size of the compact direction. It can vary smoothly over ordinary spacetime, so waves of φ are waves in the radius of the fifth dimension.

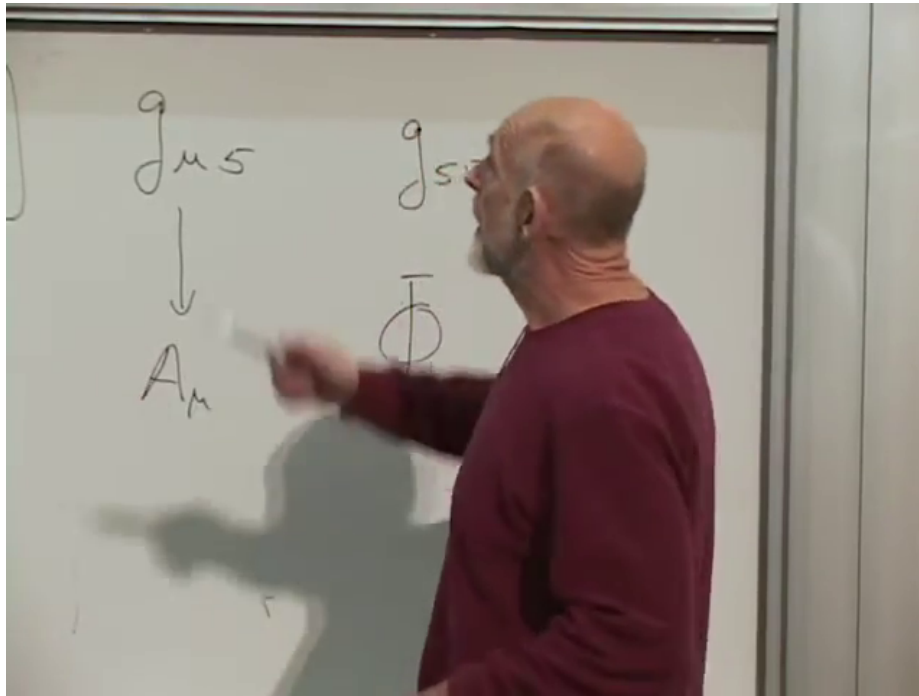


Figure 10.10: The board identifies $G_{\mu 5}$ with a vector A_μ and G_{55} with a scalar Φ .

Lecture frame: 00:47:40.

The source of a mixed gravitational component is mixed momentum flow. In particular, momentum around the fifth direction sources $G_{\mu 5}$. A four-dimensional observer therefore interprets the integer n as an electric charge coupled to A_μ . This is the Kaluza–Klein mechanism: what is momentum in the higher-dimensional geometry becomes charge in the lower-dimensional field theory.

1

Question from the audience. Is the scalar φ proportional to the momentum integer n ? ^a

Response. No. The integer n labels a quantized state and acts as charge. The scalar $\varphi(x)$ is a continuous field that determines the local compactification size. It may vary from place to place; it is not itself an integer.

^aLecture timestamp: 00:48:43.

10.7 The Second Vector Field and Winding Charge

Winding is also charge-like, but it cannot couple to the same vector in the same way. To see the second vector, return to the first closed-string excitations. A closed string has left-moving and right-moving oscillator modes. Physical states obey level matching: the total left-moving and right-moving excitation levels must agree.

In the bosonic theory the unexcited state is the tachyon. A single left-moving or right-moving oscillator would violate level matching. The first allowed massless state therefore contains one

¹Editorial clarification: The lecture loosely calls the scalar from G_{55} a “dilaton.” In this reduction it is specifically the radius modulus. The ten-dimensional string dilaton, which controls the string coupling, is a separate field, although lower-dimensional scalar combinations can mix under changes of frame and duality.

oscillator of each kind,

$$\alpha_{-1}^I \tilde{\alpha}_{-1}^J |0; k\rangle. \quad (10.15)$$

The indices I, J run over transverse spatial directions, including the compact direction. If one index is ordinary and the other is 5, there are two vector-like oscillator states,

$$\alpha_{-1}^\mu \tilde{\alpha}_{-1}^5 |0; k\rangle, \quad \alpha_{-1}^5 \tilde{\alpha}_{-1}^\mu |0; k\rangle. \quad (10.16)$$

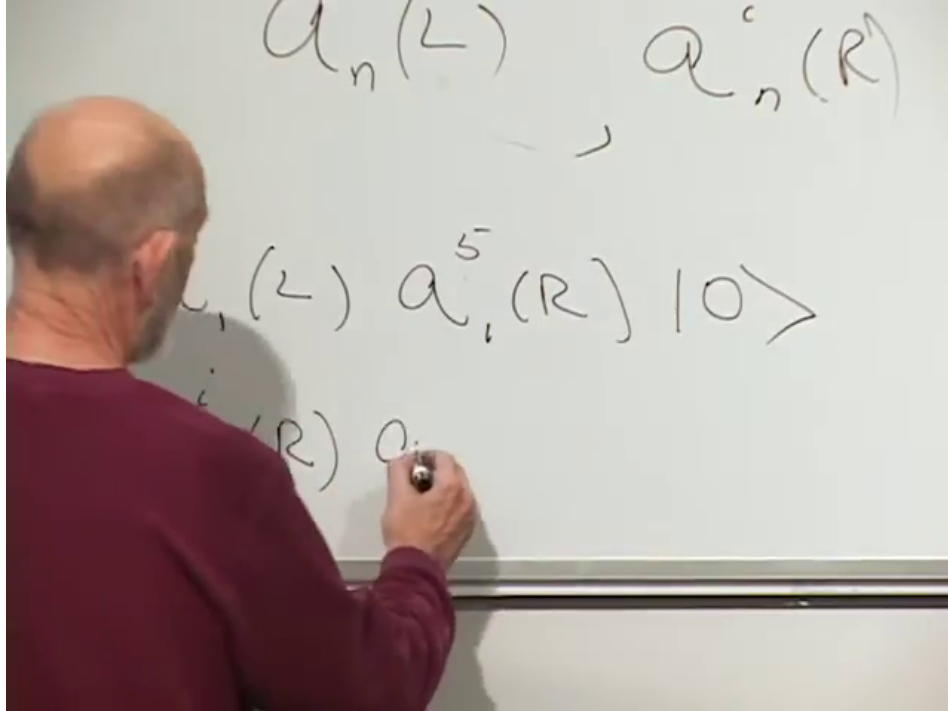


Figure 10.11: The two left–right oscillator orderings that become four-dimensional vector states when one polarization points along the compact direction.

Lecture frame: 00:58:20.

Their symmetric and antisymmetric combinations belong to different spacetime fields:

$$\alpha_{-1}^{(\mu} \tilde{\alpha}_{-1}^{5)} |0; k\rangle \longleftrightarrow G_{\mu 5}, \quad (10.17)$$

$$\alpha_{-1}^{[\mu} \tilde{\alpha}_{-1}^{5]} |0; k\rangle \longleftrightarrow B_{\mu 5}. \quad (10.18)$$

The first is the Kaluza–Klein vector sourced by compact momentum. The second descends from the antisymmetric Kalb–Ramond field and is sourced by string winding. Thus the reduced theory contains two photon-like fields, one for each charge sector.

If both oscillator indices point along the compact direction, the state is a scalar associated with the compactification modulus. If both are in ordinary transverse directions, the symmetric traceless part is the graviton, the antisymmetric part is $B_{\mu\nu}$, and the trace contributes to a scalar sector.

Question from the audience. Which of these states is the singlet and which is the triplet? ^a

Response. That classification is not the useful one here. Both mixed states transform as four-dimensional vectors. Their sum and difference separate the symmetric metric polarization from the antisymmetric B -field polarization. The state with two compact indices is instead a four-dimensional scalar.

^aLecture timestamp: 01:00:21.

The complete dictionary reached before the break is

$$\begin{array}{ll} n \leftrightarrow w, & R \leftrightarrow \frac{\alpha'}{R}, \\ \partial_\tau Y \leftrightarrow \partial_\sigma \tilde{Y}, & G_{\mu 5} \leftrightarrow B_{\mu 5}. \end{array} \quad (10.19)$$

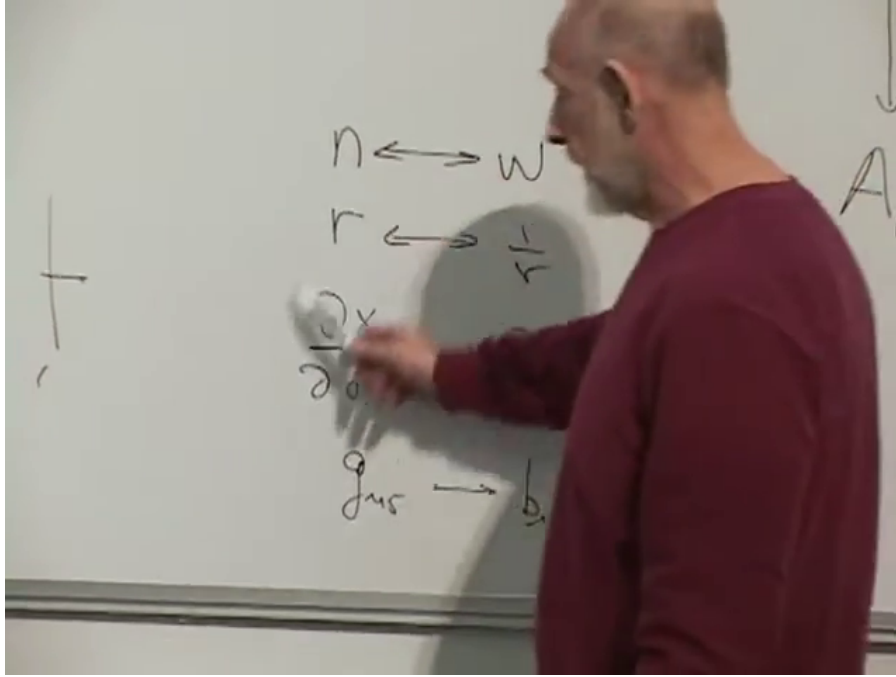


Figure 10.12: The board summary of the T-duality map: momentum and winding, radius and inverse radius, worldsheet derivatives, and the two vector fields are exchanged. The lecturer partly obscures the final line, which is therefore supplied explicitly in Equation (10.19).

Lecture frame: 01:02:20.

This construction should not be confused with a claim that every field just listed has already been identified in the observed particle spectrum. At this point it is a mathematical consequence of the compactified string theory. There are many proposals for connecting such structures to phenomenology, but that is not the problem being solved in this lecture.

10.8 Open Strings Change the Question

After the break, we begin again with open strings. A closed string can have stable winding because it has no endpoint through which the winding can slip away. An open string has endpoints and

therefore no analogous topologically protected winding number. Yet T-duality is an exact symmetry of the theory. What does the symmetry do to an open string?

For a freely moving endpoint, the variation of the string action gives a Neumann boundary condition. In the simple (X, Y) geometry,

$$\partial_\sigma X|_{\partial\Sigma} = 0, \quad \partial_\sigma Y|_{\partial\Sigma} = 0. \quad (10.20)$$

The mechanical picture is straightforward. Regard the string as a chain of ever smaller mass elements. A nonzero tangent force on the last element would give it an acceleration that diverges as its mass tends to zero. A free endpoint therefore adjusts until the net endpoint tension vanishes. Someone could instead hold the endpoint fixed and supply the necessary force, but then the endpoint would no longer be free.

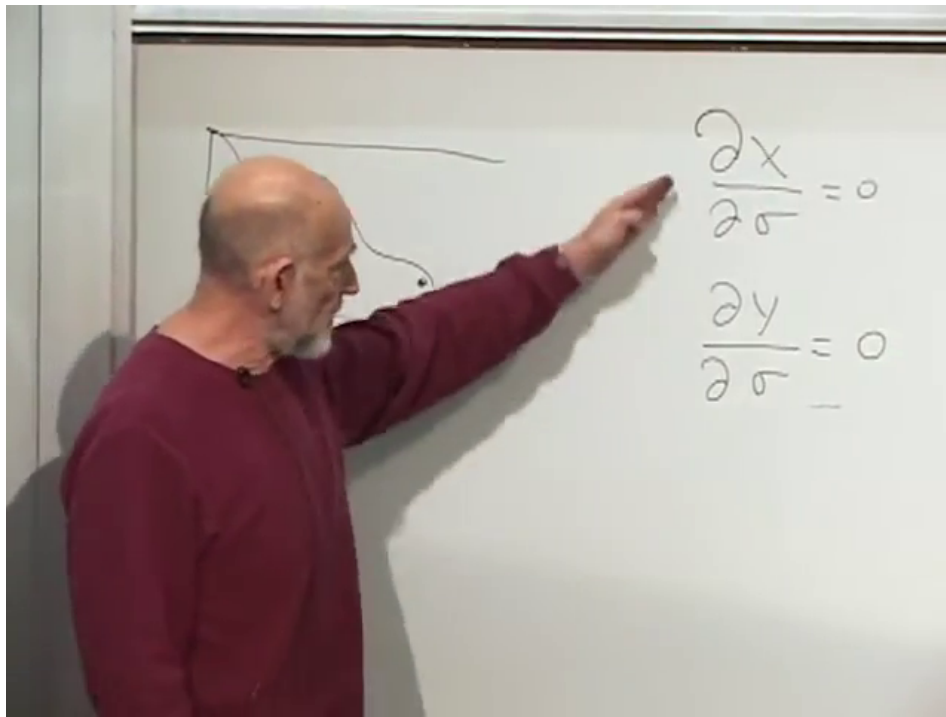


Figure 10.13: The open string and its completed Neumann endpoint conditions, $\partial_\sigma X = \partial_\sigma Y = 0$.

Lecture frame: 01:09:20.

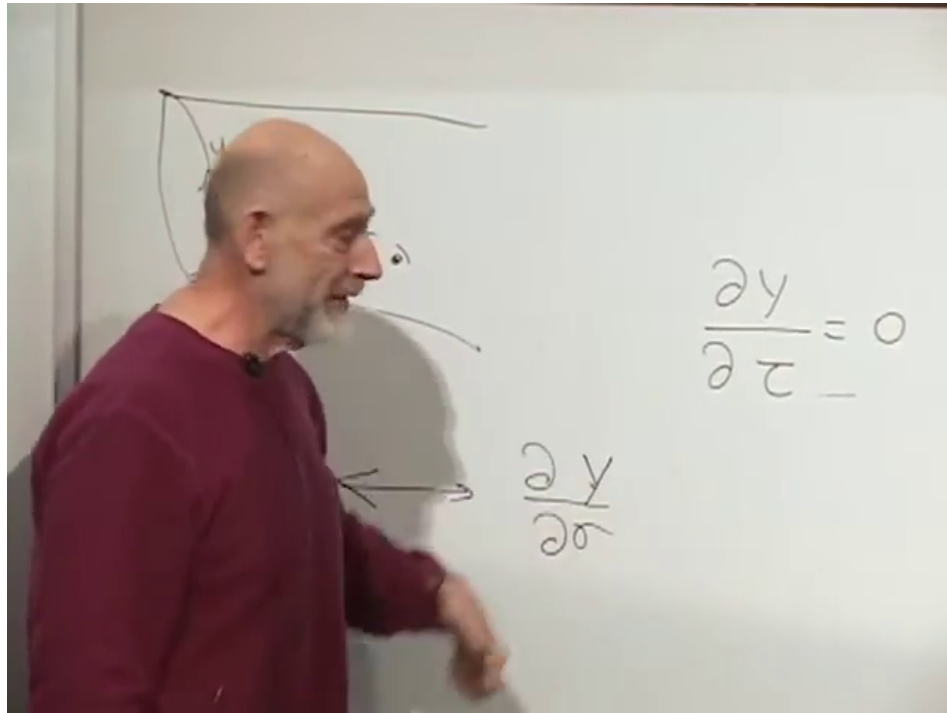
Now dualize only the compact coordinate Y . The noncompact coordinate X is unchanged. Equation (10.10) turns the second Neumann condition into

$$\partial_\tau \tilde{Y}|_{\partial\Sigma} = 0. \quad (10.21)$$

Therefore the endpoint has a time-independent position in the dual compact direction,

$$\tilde{Y}|_{\partial\Sigma} = Y_0. \quad (10.22)$$

This is a Dirichlet boundary condition.



$$\begin{array}{ccc}
 \text{Neumann} & & \text{Dirichlet} \\
 \partial_\sigma Y|_{\partial\Sigma} = 0 & \xrightarrow{T_Y} & \partial_\tau \tilde{Y}|_{\partial\Sigma} = 0 \\
 & & \tilde{Y}(\tau) = Y_0 \text{ at each endpoint}
 \end{array}$$

Figure 10.14: The completed dual endpoint condition on the board and its clean logical reconstruction. T-duality turns freedom in the compact direction into a fixed endpoint position in the dual direction.

Lecture frame: 01:12:35.

Question from the audience. Why does the dual condition mean that the endpoint is fixed?
^a

Response. Because $\partial_\tau \tilde{Y} = 0$ is zero velocity in the dualized direction. Wherever the endpoint begins in $\tilde{Y}$, it remains there. Its coordinates parallel to the un-dualized directions are still free.

^aLecture timestamp: 01:12:39.

10.9 The Object Holding the Endpoint Is a D-Brane

If an endpoint is fixed, something must be able to hold it there. Exact T-duality therefore requires a new object on which open strings can end. That object is a D-brane. The letter *D* stands for Dirichlet, not for the dimension of the object. In the historical aside of the lecture, the objects might have been called after Joseph Polchinski, whose analysis made their dynamical role unmistakable, but the already established term “p-brane” left Dirichlet as the natural name.

A *p*-brane has *p* spatial dimensions. A string is a one-brane, a membrane is a two-brane, and an extended three-dimensional object is a three-brane. A *Dp*-brane is a *p*-brane on which fundamental-

string endpoints obey Dirichlet conditions in the transverse directions and Neumann conditions in the directions tangent to the brane.

Nothing in the duality argument singles out one value of Y_0 , so the D-brane cannot be a rigid fixture nailed to a preferred point in the universe. It must be movable. Relativity also forbids an absolutely rigid extended body, so a D-brane can fluctuate and bend. The simple flat brane drawn on the board is its equilibrium configuration, not a restriction on its dynamics.

Question from the audience. Must the two endpoints of an open string have the same fixed value of the dual coordinate? ^a

Response. Not in general. Both endpoints may lie on the same D-brane, in which case they share that brane's transverse location. They may also lie on two different D-branes at different transverse positions, and the string then stretches between them. A brane itself may fluctuate and bend, so the endpoint locus need not remain a perfectly flat plane.

^aLecture timestamp: 01:14:55.

10.10 Counting the Dimensions of a D-Brane

Ten-dimensional superstring theory has nine spatial dimensions. Suppose T-duality imposes a Dirichlet condition in one spatial direction. The endpoint is still free in the other eight, so its allowed locus is an eight-dimensional D8-brane. Each additional independent Dirichlet condition removes one direction of endpoint motion. If k spatial coordinates are fixed, then

$$p = 9 - k. \quad (10.23)$$

The tabletop gives the three-dimensional version of the same count. Fix the vertical coordinate and a fingertip may move on a two-dimensional tabletop. Fix a second coordinate and it may move only along a line. Fix all three and it is confined to a point.

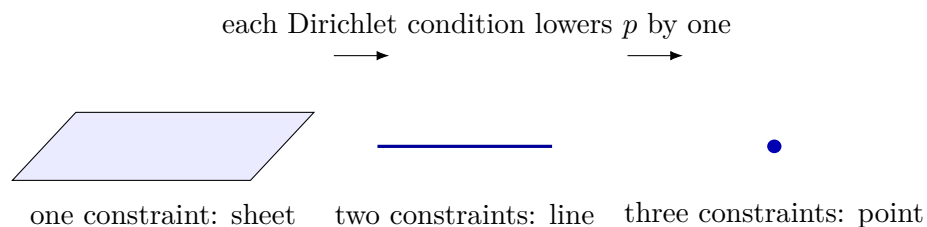


Figure 10.15: The dimensional-counting argument used at the tabletop. A Dp -brane is the locus in which endpoints remain free to move.

Repeated dualities generate D8-, D7-, D6-branes and so on down to a D0-brane, which is pointlike in space. If no spatial direction is fixed, a D9-brane fills all nine spatial dimensions and ordinary open strings are free to move everywhere. The lecture briefly mentions $D(-1)$ -branes; these are localized in Euclidean spacetime and behave as instantons rather than as ordinary particles propagating through time.

²Editorial clarification: The dimensional ladder is a statement about the type-II theories taken together. Stable

Question from the audience. If only six dimensions are compact, should the counting start by subtracting from six rather than nine? ^a

Response. The elementary derivation performs T-duality on compact directions, but a dual circle becomes arbitrarily large as the original circle becomes arbitrarily small. In that limit it is effectively noncompact. The resulting branes can therefore extend along or be transverse to directions that appear noncompact at accessible scales. The dimension p counts all spatial directions tangent to the brane, not only the compact ones used in one derivation.

^aLecture timestamp: 01:20:49.

A D0-brane is a new kind of particle. An open string may run from one D0-brane to another, return to the same D0-brane, or stretch toward a very distant brane. A string taken literally to spatial infinity has infinite length and hence infinite energy, so the physically useful version places its other endpoint on another object.

The next case, a D1-brane, is itself string-shaped, but it is not the fundamental string with which we began. It can bend and close on itself, and fundamental strings may end on it. Its tension scales differently with the string coupling:

$$T_{D1} = \frac{1}{2\pi\alpha'g_s}, \quad T_{F1} = \frac{1}{2\pi\alpha'}. \quad (10.24)$$

At weak coupling $g_s \ll 1$, the D-string is much heavier than the fundamental string, which explains why it first appears as a massive anchor.

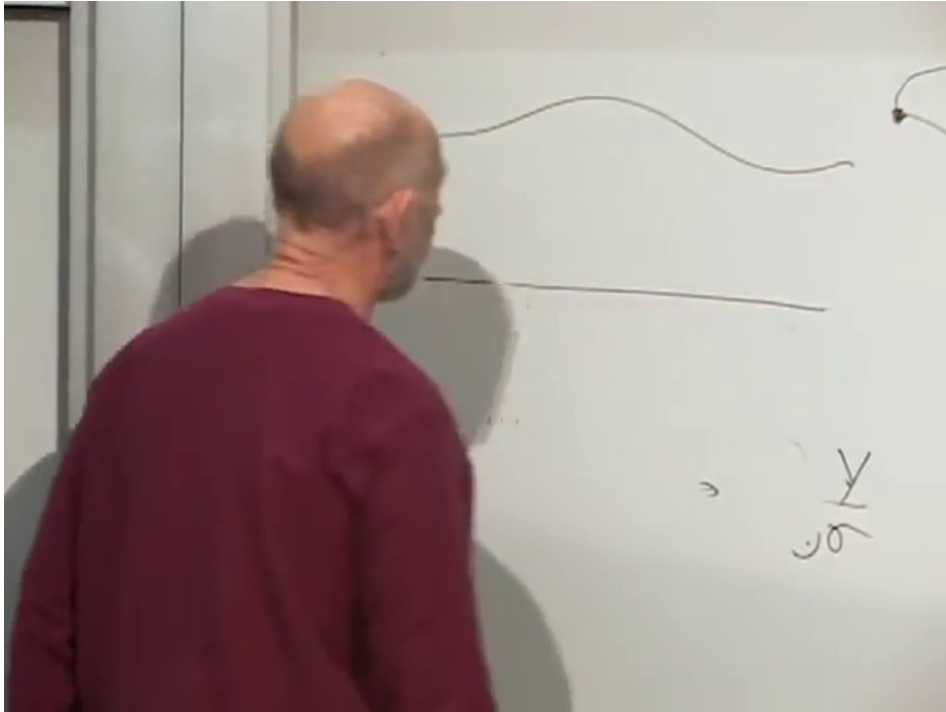


Figure 10.16: The board's D1-brane: a bendable one-dimensional object, distinct from the fundamental strings that may end on it.

Lecture frame: 01:24:20.

BPS Dp -branes have even p in type IIA and odd p in type IIB. T-duality along one circle changes p by one and simultaneously exchanges type IIA with type IIB.

Question from the audience. Does T-duality act on the D-branes as well as on the strings?
^a

Response. Yes. A T-duality parallel to a Dp -brane lowers its dimension to $D(p - 1)$, while a T-duality transverse to it raises the dimension to $D(p + 1)$. Branes are part of the duality dictionary, not external supports left unchanged by it.

^aLecture timestamp: 01:25:35.

Question from the audience. May directions tangent to a brane still be compact? ^a

Response. Yes. “Tangent” and “compact” answer different questions. A direction may lie along the brane and still close into a circle; it may also be noncompact. What matters for the endpoint is whether the boundary condition in that direction is Neumann or Dirichlet.

^aLecture timestamp: 01:27:25.

A D2-brane is a membrane-like sheet. A D3-brane has three spatial dimensions. In a genuinely higher-dimensional spacetime a D3-brane is a three-dimensional subspace rather than all of space, even though observers confined to it may experience it as their entire spatial world.

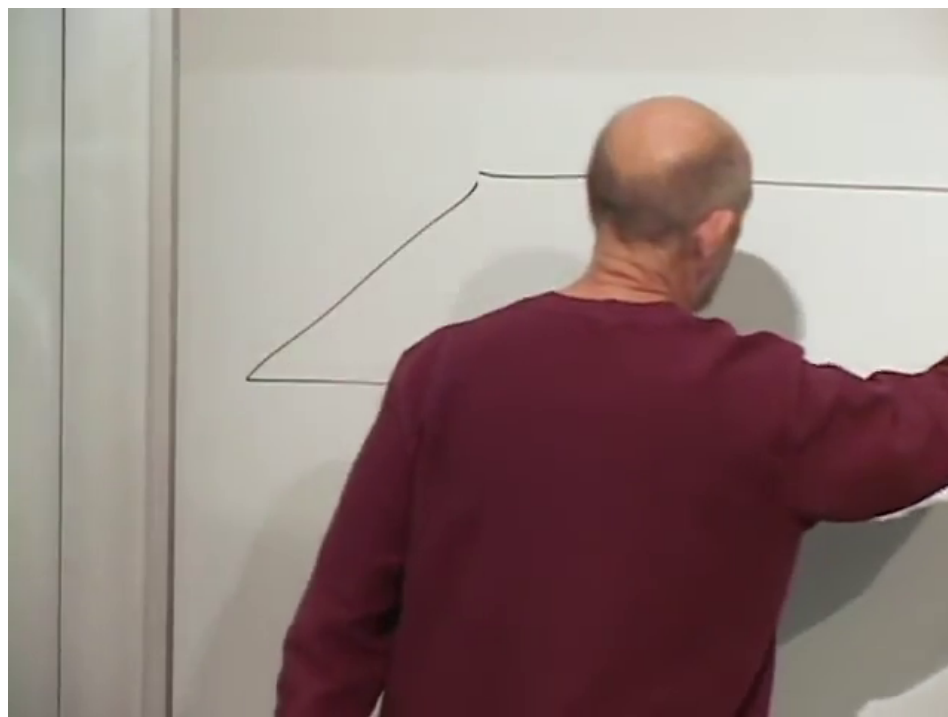


Figure 10.17: A D2-brane drawn as a sheet on which string endpoints can move.

Lecture frame: 01:27:55.

The derivation given here is intentionally a sketch. It is not by itself a proof that string theory describes nature. Within string theory, however, the consistency checks are numerous: dualities, quantum calculations, and mathematical results all point to D-branes as genuine dynamical objects.

10.11 Open Strings Become Fields on a Brane

Take an otherwise empty D2-brane and place short open strings on it. Each endpoint is confined to the sheet, so the whole low-energy excitation moves along the sheet. Two oriented strings whose adjacent endpoints carry compatible arrows can join. The joined string may later split again. Read as particle histories, these are the vertices of Feynman diagrams: two particles enter, an intermediate excitation propagates, and two particles emerge.

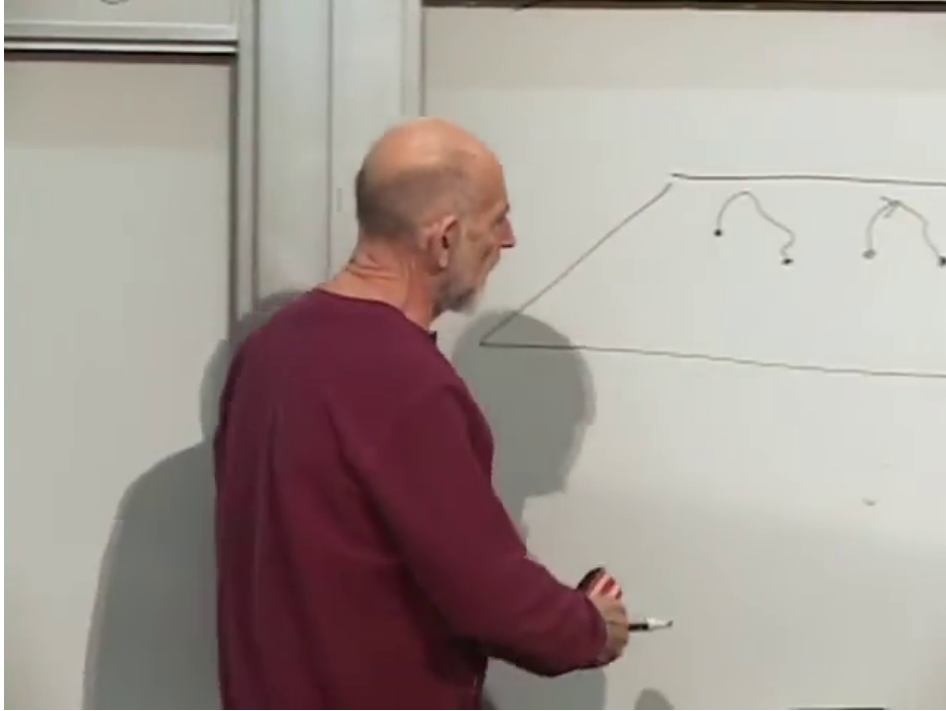


Figure 10.18: Oriented open strings on a D2-brane joining through compatible endpoints. At low energy the process is read as a field-theory interaction on the brane.

Lecture frame: 01:29:55.

At energies well below the string scale, the internal vibrational tower is not excited. The remaining massless open-string state is a vector field on the brane worldvolume. For one brane this is an abelian $U(1)$ gauge field, the brane analogue of electromagnetism. The field does not propagate through the full ten-dimensional bulk; its quanta are open strings whose endpoints keep them on the brane.

10.12 A Stack of Branes Produces Nonabelian Color

Now place several parallel branes near one another. For three branes, label them Red, Green, and Blue. An oriented string has an ordered pair of endpoint labels: it may run from Red to Green, Green to Blue, Red to Red, and so on. With N branes the massless open-string states can be arranged as an $N \times N$ matrix

$$A_{\mu}^i{}_j, \quad i, j = 1, \dots, N. \quad (10.25)$$

These are the Chan–Paton labels of a $U(N)$ gauge theory.

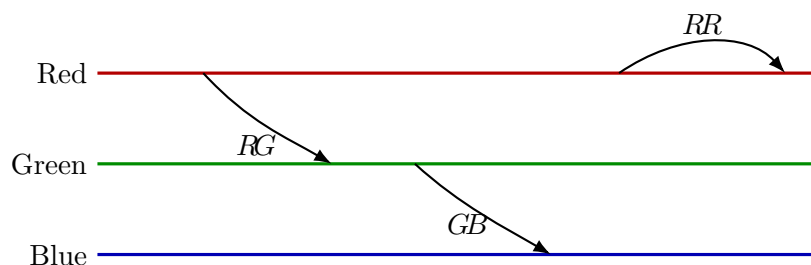
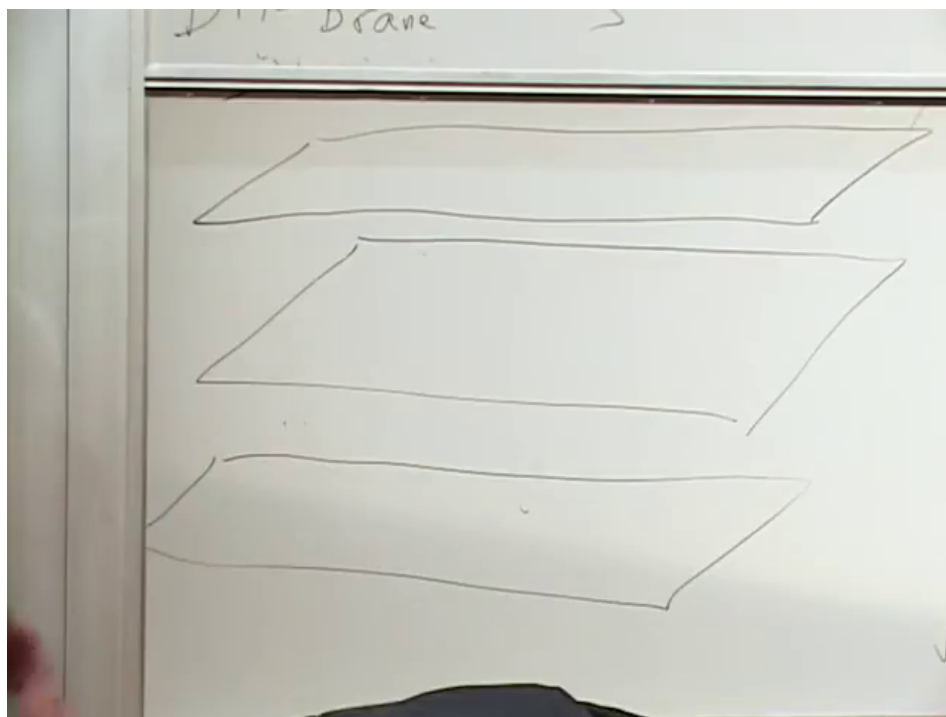


Figure 10.19: Three parallel branes on the board and a clean reconstruction of the ordered endpoint labels carried by strings between them.

Lecture frame: 01:33:15.

Question from the audience. Can a string really run from the Red brane to the Green brane without being blocked by the other drawings? ^a

Response. Yes. The branes are embedded in a higher-dimensional space. The two-dimensional board picture is only a projection, so a string may pass around another projected line in a transverse direction.

^aLecture timestamp: 01:35:07.

Question from the audience. Must a Red–Green string combine with a Green–Blue string in order to become Red–Blue? ^a

Response. That is the basic local joining rule. Compatible endpoints meet, and the two segments become one string whose unjoined endpoints retain their Red and Blue labels. More elaborate processes are built by repeating the same operation.

^aLecture timestamp: 01:35:39.

The joining rule is matrix multiplication made geometric. A Red–Green string can join a Green–Blue string because the Green endpoint of one is compatible with the Green endpoint of the other. The result is a Red–Blue string:

$$(R, G) \circ (G, B) \longrightarrow (R, B). \quad (10.26)$$

The repeated index is joined and disappears internally. This is the same index structure that appears at a Yang–Mills interaction vertex.

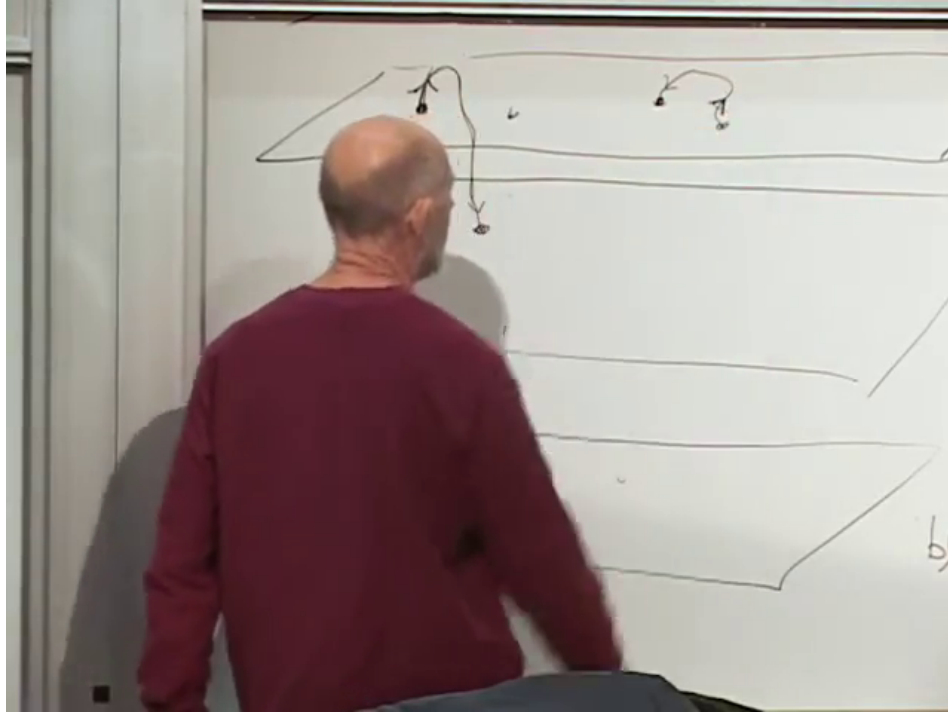


Figure 10.20: Joining open strings on a stack. Endpoint orientation enforces the same ordered-index composition used by nonabelian gauge fields.

Lecture frame: 01:36:30.

For three coincident branes the complete open-string gauge group is $U(3)$, with nine vector states. Its algebra decomposes as

$$\mathfrak{u}(3) = \mathfrak{su}(3) \oplus \mathfrak{u}(1). \quad (10.27)$$

The eight traceless combinations have the color structure of the QCD gluons, while the diagonal center-of-mass $U(1)$ is a separate abelian field.

3

Quark-like matter carries only one color index. In the brane picture it is represented by a string with one endpoint on the color stack and the other on a different, distant brane. Relative to the color stack, only one end is visible as a color label. Reversing the orientation gives the antiparticle.

³Editorial clarification: The lecture describes one of the nine color-pair combinations as able to detach and disappear, thereby leaving eight gluons. The precise D-brane statement is that a stack of three coincident branes gives $U(3)$, not only $SU(3)$. Obtaining the QCD octet requires separating, projecting out, or otherwise decoupling the diagonal $U(1)$. The lecture's picture correctly motivates the ordered color indices, but it is not a complete derivation of QCD.

A quark-like and antiquark-like string with compatible endpoints can join and leave the color stack; two identically oriented quark-like strings cannot annihilate in that way.

The resulting low-energy theories are generally supersymmetric Yang–Mills theories, with additional matter determined by the branes and their intersections. They are close enough to QCD to illuminate nonabelian dynamics, but they are not automatically the Standard Model. With only one brane, the same construction reduces to a $U(1)$ theory: the short brane–brane strings are photon-like, while oriented strings reaching a second sector play the roles of positive and negative electric charges.

10.13 D-Strings End as Magnetic Monopoles

There are now two kinds of string-like object in the story. The original fundamental string, or F-string, carries electric flux on a D3-brane. The D1-brane, or D-string, is much heavier at weak coupling. A D-string can also end on a D3-brane. Its endpoint is not electrically charged under the worldvolume gauge field; it is magnetically charged. From the perspective of observers on the D3-brane, the endpoint is a magnetic monopole.

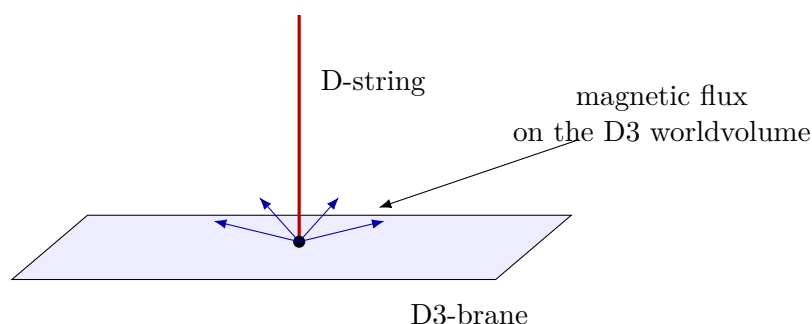
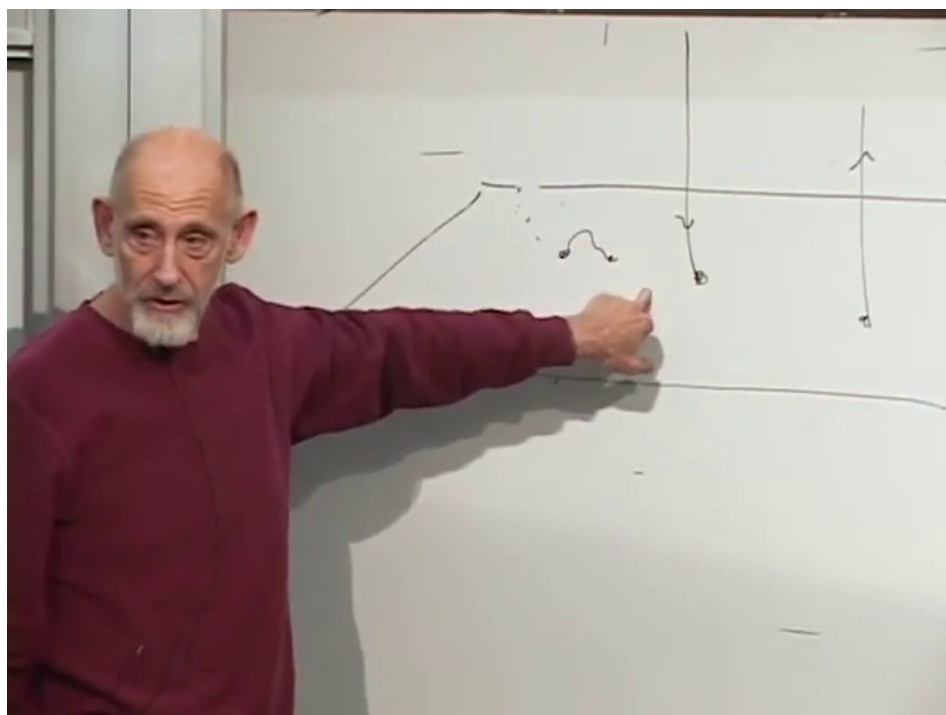


Figure 10.21: The board endpoint construction and its clean interpretation. A D-string ending on a D3-brane appears as a magnetic monopole to the worldvolume gauge theory.

Lecture frame: 01:41:30.

The mass hierarchy also has the right qualitative form. Since the D-string tension carries a factor $1/g_s$, its endpoint is much heavier than an ordinary electric excitation at weak coupling. The electric/magnetic relation between F-string and D-string endpoints is part of the wider duality structure of type IIB string theory.

Question from the audience. Is the endpoint of the D-string neutral? ^a

Response. It is neutral with respect to the electric charge carried by a fundamental-string endpoint, but it is not uncharged. It carries magnetic charge and is therefore a monopole in the D3-brane gauge theory.

^aLecture timestamp: 01:42:19.

10.14 What the Construction Does and Does Not Claim

The path from T-duality to branes was assembled over decades, not in a single calculation. A theory first proposed to describe hadrons became a candidate theory of quantum gravity. Its dualities then forced D-branes into the spectrum, and gauge theories reappeared as the low-energy dynamics of strings ending on those branes. In that sense the subject made a full circle.

D-brane and related dual constructions have become powerful tools for studying supersymmetric gauge theories, strongly coupled systems, and connections between gravity and field theory. They have also informed models of fluids and plasmas through gauge/gravity duality. This success is not the same as deriving the observed particle spectrum directly. The compactification geometry, brane configuration, supersymmetry breaking, and many other ingredients still have to be selected.

Question from the audience. What determines which compact dimension is dualized? ^a

Response. Mathematically, one may dualize any compact direction for which the required symmetry is present, choosing the description that is most convenient. Physically, the brane orientations and compactification data realized in a particular state are selected by the history and dynamics of that state.

^aLecture timestamp: 01:44:59.

Question from the audience. Which brane supports the monopole construction? ^a

Response. The world experienced by the gauge theory is a D3-brane. The D-string extends into a transverse direction and ends on that D3-brane; its endpoint is the monopole seen in the three spatial dimensions of the brane.

^aLecture timestamp: 01:45:21.

Question from the audience. Are D-strings oriented as well? ^a

Response. Yes. Reversing a D-string's orientation reverses the corresponding charge. The two endpoint orientations distinguish monopole from antimonopole just as orientation distinguishes electric charge for the fundamental-string endpoints.

^aLecture timestamp: 01:45:55.

Question from the audience. Can this simple toroidal picture account for the left–right asymmetry of chiral particle physics? ^a

Response. Not by itself. A flat torus is too symmetric. Chiral spectra require more structure, such as asymmetric brane intersections, orbifold projections, or compactification on suitable Calabi–Yau geometries. String theory has mechanisms for chirality, but they lie beyond this elementary T-duality construction.

^aLecture timestamp: 01:46:05.

We have followed one symmetry through four concrete forms: an exchange of spectra, an exchange of worldsheet derivatives, an exchange of spacetime fields, and finally an exchange of endpoint boundary conditions. The last form is the decisive one. It turns a mathematical equivalence into the discovery of D-branes, and it turns the dynamics of open strings into gauge theory on their worldvolumes. At that point, in Susskind's phrase, the lecture has used up its time, energy, and momentum.