

MATRIX METHODS

Network parameter characterisation — S-, Z- and Y-parameters for RF and high-speed digital design

September 2026 • Prepared for Brendan Lynskey

What this report is about, in plain terms. Consider a system with N ports — a filter, a length of cable, an amplifier, or simply a box with N coax connectors on it. Because such a system is linear, its behaviour at any one frequency is captured entirely by an $N \times N$ matrix, and which matrix you get depends only on what you choose to call the input. If you drive the ports with currents and read off the voltages that appear, the matrix is the impedance matrix Z , measured in ohms. Turn that around — drive with voltages and read the currents — and you have the admittance matrix Y , in amps per volt. Work instead with the travelling waves that a transmission line naturally carries, asking how the wave leaving one port compares with the wave arriving at another, and you have the scattering matrix S . All three are functions of frequency, so a single spot measurement and a full swept characterisation are not different ideas: the sweep is the same matrix evaluated at many frequencies instead of one. What genuinely is a second story is the time domain, where the network's internal dynamics live and where the shape of a swept curve comes from. This is a companion article to 'Matrix Concepts in Digital Filter Design'.

If you want...	Turn to...	What you will find
A quick map of the whole report	§1	Which matrix property answers which laboratory question, and the three different boundaries that all happen to look like $ \cdot = 1$
To know what an S-parameter actually is	§4.2	S_{ij} defined as the relative magnitude and phase from port j to port i , with S_{21} and S_{11} spelled out
To know why S-parameters exist at all	§4.1	The transmission-line argument, plus an honest note on what a scattering measurement cannot tell you
Voltage waves or power waves?	§8.1	Why the two agree at a real Z_0 and genuinely differ at a complex one
To know why a response peaks and dips	§5	Poles and zeros as eigenvalues of the internal state matrix, with a worked example in numbers
To know why fitting measured data is subtle	§6	Causality is easy to impose, passivity is not; the Hamiltonian test
The Smith chart without the mystery	§8	The one-port case of the same matrix Möbius map
Differential pairs done properly	§9	Mixed-mode as an orthogonal change of basis; the four quadrants; how much mode conversion a given skew buys you
A checklist for data from the bench	§11	Reciprocity, causality, passivity and de-embedding, in the order worth checking them

Notation. Bold lower-case symbols ($\mathbf{V}$, $\mathbf{I}$, $\mathbf{a}$, $\mathbf{b}$) are vectors holding one entry per port, and upper-case symbols (Z , Y , S) are the $N \times N$ matrices that relate them. M^H denotes the conjugate transpose and M^T the plain transpose; I is the identity; the symbol ≥ 0 means positive semidefinite. The complex frequency variable is $s = \sigma + j\omega$, and the imaginary axis $s = j\omega$ is both the stability boundary and the axis along which a frequency response is plotted.

Unless stated otherwise the S-parameter reference impedance Z_0 is a single real value, 50 Ω , common to every port.

1 Two questions, and why only one of them is new

Almost everything you ask of a multiport network falls into one of two categories, and it is worth separating them at the outset because the previous edition of this report ran them together.

The first category is about the network seen from outside. How much of the signal incident on port 1 reaches port 2, and with what phase shift? Does the network behave the same way when you swap source and receiver? Could it be passive, or is it quietly supplying power? Every question of this sort is answered by the entries of $S(f)$, $Z(f)$ or $Y(f)$ and by the algebraic structure those matrices carry — symmetry, definiteness, unitarity. It makes no difference whether you ask at one frequency or across a two-decade sweep. The matrix is a function of frequency, and a swept measurement is simply that function sampled at many points rather than one. There is a single concept here, evaluated once or evaluated a thousand times, and the frequency response is what you get when you plot one of its entries.

The second category is about what is happening inside. Why does $|S_{21}|$ climb to a peak near 1 GHz and then collapse into a deep notch at 2 GHz? That shape is not arbitrary, and it is not explained by anything in the matrix at a single frequency. It comes from the network's internal dynamics — from the differential equations obeyed by the capacitor voltages and inductor currents — and those dynamics are most naturally described in the time domain, as an impulse response. The frequency response and the impulse response are a Laplace pair, two faces of the same network, one plotted against frequency and the other against time. The matrix that governs the time-domain side is the state matrix A_c , and its eigenvalues are the poles that set the shape of the swept curve.

So when this report contrasts two pictures, it means the frequency response and the impulse response, not one frequency against many. Sections 2 to 4 stay entirely in the frequency domain. Section 5 crosses to the time domain and shows how the two views are tied together.

Three boundaries that all look like $|\cdot| = 1$. The condition that some quantity has unit modulus turns up three times in this material, and it means something different each time. **(1) The eigenvalues of $S(f)$.** At a fixed frequency, S is an $N \times N$ matrix with eigenvalues λ . If the network is lossless then S is unitary and every $|\lambda| = 1$. This is a statement about how the ports exchange energy at that frequency; move to another frequency and you get another set of eigenvalues. Nothing about stability is implied. **(2) The poles of $S_{21}(s)$ in the s -plane.** These are the eigenvalues of A_c . A continuous-time network is stable when they all lie strictly to the left of the $j\omega$ -axis, and that same axis is the line along which the frequency response is evaluated. **(3) The poles of $H(z)$ in the z -plane,** which belong to the discrete-time world of the companion article. There the boundary is the circle $|z| = 1$: poles inside it are stable, and the response is evaluated on the circle itself. The pair worth keeping apart is (1) and (3), because both are unit circles drawn in a complex plane and neither has anything to do with the other.

2 The N-port network: one system, three descriptions

A port is a pair of terminals, or the reference plane of a transmission line, at which energy enters or leaves. Each port i has a voltage V_i across it and a current I_i flowing into it, and by convention the currents are always defined as flowing inwards so that the sign bookkeeping in the power expressions comes out cleanly. Gather those quantities into vectors $\mathbf{V}$ and $\mathbf{I}$ of length N .

Because the network is linear, the port voltages and the port currents are related by a single matrix, and which of the two you treat as the independent variable is entirely your choice. Take the currents as the input and you obtain the impedance matrix, $\mathbf{V} = \mathbf{Z} \mathbf{I}$, whose entries have units of ohms. Take the voltages instead and you obtain the admittance matrix, $\mathbf{I} = \mathbf{Y} \mathbf{V}$, whose entries are in siemens — amps per volt — and which is simply the inverse of the first, $\mathbf{Y} = \mathbf{Z}^{-1}$, whenever that inverse exists. Abandon voltages and currents altogether in favour of the travelling waves discussed in §4, and you obtain the scattering matrix $\mathbf{S}$, whose entries are dimensionless ratios of one wave amplitude to another. The three matrices are three coordinate systems laid over the same underlying linear operator; §8 gives the formulae for moving between them.

All three depend on frequency. Writing $\mathbf{Z}$ rather than $\mathbf{Z}(f)$ is only a shorthand, and nothing in the algebra that follows is confined to a single frequency: apply it at every point of a sweep and you have characterised the network over a band.

Port conventions: currents defined into the network

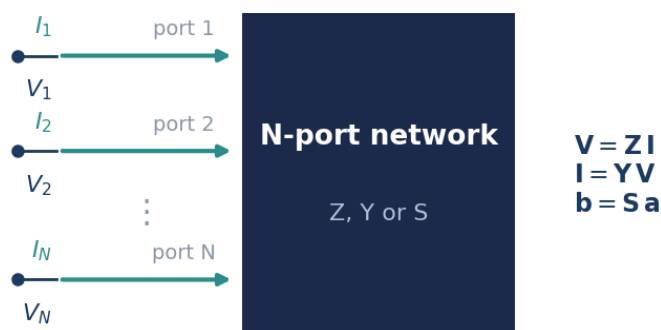


Figure 1 — Port conventions. Currents are defined into the network, and $\mathbf{Z}$, $\mathbf{Y}$ and $\mathbf{S}$ are three ways of describing the same linear relationship between what goes in and what comes out.

3 $\mathbf{Z}$ and $\mathbf{Y}$: symmetry means reciprocity, definiteness means passivity

3.1 What the two matrices mean on the bench

Column j of the impedance matrix has a direct experimental meaning. Inject a known current into port j , leave every other port open-circuited so that no current can flow there, and record the voltage that appears at each port; the ratio V_i / I_j is the entry Z_{ij} . The admittance matrix is the mirror image: drive port j with a known voltage, short-circuit every other port, and record the currents, so that $Y_{ij} = I_i / V_j$.

At audio and low radio frequencies this is a practical recipe, and it becomes more demanding as frequency rises, because an open circuit has to be an open circuit at the reference plane. A physical open has fringing capacitance and may radiate; a physical short has series inductance. It is worth being precise about how much of a problem that really is, because it is often overstated. Open, short and load standards are exactly what a conventional SOLT calibration uses, routinely and accurately, at 5 GHz and a good deal higher — the standards are not assumed to be ideal but are characterised, their fringing capacitance and residual inductance captured in a polynomial model that the analyser applies. Higher still, and on wafer, schemes such as TRL take over because they lean on transmission-line sections rather than on lumped terminations. So realising ideal terminations is challenging rather than hopeless, and the difficulty is a real inconvenience but not, by itself, the reason the industry works in S-parameters. Section 4.1 gives the better reason.

3.2 Reciprocity: swapping source and probe changes nothing

If the network is built from ordinary resistors, capacitors, inductors and PCB copper, then transmission from port 1 to port 2 is identical to transmission from port 2 to port 1. The matrix records this as symmetry about its diagonal, $Z = Z^T$, and equivalently $Y = Y^T$ and $S = S^T$. Note that this is the plain transpose and not the conjugate transpose: Z is complex, and reciprocity is a weaker statement than Hermitian symmetry. Ferrite isolators, circulators and active devices break reciprocity deliberately, which is precisely what makes them useful, and in a measurement the size of the antisymmetric part $Z - Z^T$ is a useful sanity check on data that ought to be reciprocal.

3.3 Passivity: the network cannot manufacture power

A passive network can absorb energy and can store it, but it cannot on average deliver more than it receives, no matter how you drive it. Expressed in the port variables, the average power delivered into the network by a set of phasor drives is $P_{\text{avg}} = \frac{1}{2} \text{Re}(\mathbf{V}^H \mathbf{I}) = \frac{1}{2} \mathbf{I}^H \mathbf{Z}_H \mathbf{I}$, where $\mathbf{Z}_H = (\mathbf{Z} + \mathbf{Z}^H)/2$ is the Hermitian part of the impedance matrix. Requiring that this be non-negative for every possible drive vector is exactly the statement that $\mathbf{Z}_H$ is positive semidefinite, so passivity reduces to a real eigenvalue test: every eigenvalue of $\mathbf{Z}_H$ must be at least zero. The same test applies to $\mathbf{Y}_H$. Splitting Z into its real and imaginary parts as $Z = R + jX$ makes the physical reading obvious — only R absorbs average power, X merely stores and returns it, and a lossless network is the boundary case $R = 0$.

Symmetric, Hermitian and real are three different things. This trips people up often enough to be worth stating carefully. Transpose symmetry, $Z = Z^T$, expresses reciprocity and requires a reciprocal medium; it makes no assumption about the signals being real, and Z stays complex throughout. The Hermitian condition $\mathbf{Z}_H \geq 0$ expresses passivity and likewise assumes nothing about real signals, because average power is always a Hermitian quadratic form even for complex phasors. What genuinely does require real time-domain signals is conjugate symmetry — the rule that $H(-j\omega) = H^*(j\omega)$, that the real part is even and the imaginary part odd, that $S(-f) = S^*(f)$. A vector network analyser measuring real voltages obeys it; a complex-baseband or IQ model in simulation generally does not. The practical consequence is that you may trust the Hermitian passivity test even for complex signals, but you should distrust any argument that appeals to a negative-frequency mirror unless the impulse response has been stated to be real.

4 The S matrix: travelling waves and unitary structure

4.1 Why travelling waves are the natural variables

Once a connection is long enough that the signal's phase changes appreciably along it, the natural solution of the equations describing that connection is not a single voltage but a pair of waves, one heading towards the load and one heading back. Transmission-line theory is built on that pair from the beginning, and the quantities an RF engineer works with day to day — return loss, insertion loss, VSWR, group delay, the action of a matching network, the effect of a length of line — are all most naturally written as relationships between them.

Scattering parameters carry that description into network theory, and that is the real reason they exist. When a signal arrives at one port and leaves by another, the useful thing to know is how its magnitude and phase have been altered on the way, and an S-parameter is precisely that number. Cascade two components and the phases add; insert a quarter-wave line and you can read what it does directly. The description composes the way the hardware does, which is what makes it the natural bookkeeping for a system in which transmission-line behaviour matters.

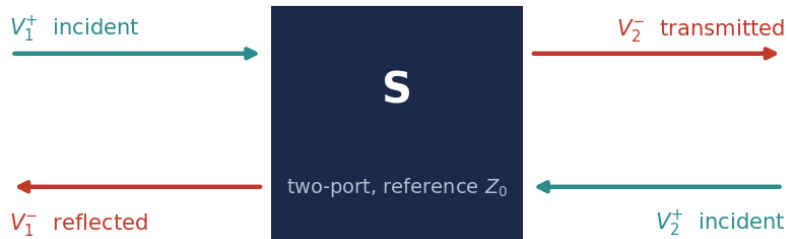
There is a secondary and quite genuine practical benefit. Because S is defined with every port that is not being driven terminated in the reference impedance rather than opened or

shorted, the measurement keeps the device in the same broadband-matched state it will meet in the real system, and a good $50\ \Omega$ load is easier to realise and to hold over a wide sweep than an ideal open or short. That is a convenience worth having, but it is a consequence of the choice of variables rather than the motivation for it.

4.2 What an S-parameter actually is

Start with what is physically on the line. At a port on a transmission line of characteristic impedance Z_0 , the total voltage and the total current are the superposition of a wave travelling towards the network and one travelling back out of it: $V = V^+ + V^-$ and $I = (V^+ - V^-)/Z_0$. Solving for the two waves gives $V^+ = (V + Z_0 I)/2$ and $V^- = (V - Z_0 I)/2$. These are voltages, measured in volts. They are what a time-domain reflectometer displays, and what superposes at an impedance discontinuity.

A scattering parameter is a ratio of one of these to another: $S_{ij} = V_i^- / V_j^+$, evaluated with $V_k^+ = 0$ for every $k \neq j$ — that is, with every other port terminated in Z_0 so that nothing is incident there. In words, S_{ij} is the complex ratio of the wave leaving port i to the wave arriving at port j , and being complex it carries a magnitude and a phase. So S_{21} is the magnitude and phase of the signal emerging at port 2 relative to the signal driven into port 1, at whichever frequency you happen to be looking at; sweep the source across a band and the collection of those numbers is the transmission response of the network. S_{11} is the wave that comes back out of port 1 relative to the one you sent in, which is the input reflection coefficient. In general the diagonal entries are reflections and the off-diagonal entries are transmissions, and the second subscript is always the port you drove.



$$S_{21} = V_2^- / V_1^+ \text{ with } V_2^+ = 0 \quad \text{— magnitude and phase at port 2 relative to the wave driven into port 1}$$

Figure 2 — The two-port wave picture. S maps the incident amplitudes onto the emerging ones, and each entry records both how much of the signal survives and how much its phase has shifted.

It is often convenient to rescale these amplitudes as $a_i = V_i^+ / \sqrt{Z_0}$ and $b_i = V_i^- / \sqrt{Z_0}$, so that $|a_i|^2$ is the power flowing into port i and $|b_i|^2$ the power flowing out of it, and the scattering matrix becomes the map $\mathbf{b} = \mathbf{S} \mathbf{a}$. The rescaling is by a constant common to every port, so it cancels from every ratio and leaves each S_{ij} exactly as it was; its only purpose is to let the power bookkeeping of §4.3 be written without carrying factors of $1/2Z_0$ around. With Z_0 real these normalised amplitudes are also Kurokawa's power waves. The two descriptions part company only when the reference impedance is complex, which §8.1 returns to.

Reading the numbers on the bench. Send 1 W into port 1 of a two-port with the other port matched. A fraction $|S_{11}|^2$ comes straight back and a fraction $|S_{21}|^2$ arrives at port 2, so $|S_{11}| = -10$ dB means a tenth of the power is reflected, and $|S_{21}| = -3$ dB means about half of it gets through. A matched 6 dB attenuator is very nearly $\begin{bmatrix} 0 & 0.5 \\ 0.5 & 0 \end{bmatrix}$. One trap is worth flagging: cascading two components is not a matter of multiplying their S matrices, because the wave leaving the first is partly reflected by the second and returns to be re-reflected. Convert to ABCD or T parameters, multiply there, and convert back. A second trap is the decibel convention. S_{21} is a ratio of wave amplitudes, so $|S_{21}|$ is a voltage ratio while $|S_{21}|^2$ is the power ratio, and the figure quoted on every datasheet is $20 \log|S_{21}|$ rather than $10 \log$. Halving the voltage is -6 dB and quarters the power; -3 dB is $|S_{21}| = 0.707$, which halves it.

What a scattering measurement does not tell you. S-parameters account for power that leaves through a port. They say nothing about where the rest of it went. A perfect open circuit reflects everything and does so in phase, sitting at $S_{11} = +1$; but a structure that reflects less than everything may be dissipating the difference in copper and dielectric, or radiating it into the room as an antenna, and no scattering measurement can separate the two. From a complete N-port set you can compute the total power that failed to emerge, $1 - \sum_i |S_{i1}|^2$, but that single number lumps dissipation and radiation together. Distinguishing them needs a different experiment — a calorimetric one, or a pattern measurement in an anechoic chamber. The passivity condition of §4.3 has the same limitation: it certifies that the network is not supplying energy, not which mechanism is consuming it.

4.3 Losslessness and passivity: unitary matrices and contractions

Count the power. The difference between what goes in and what comes out is $\frac{1}{2}(\mathbf{a}^H \mathbf{a} - \mathbf{b}^H \mathbf{b}) = \frac{1}{2} \mathbf{a}^H (\mathbf{I} - S^H S) \mathbf{a}$. If the network absorbs nothing whatever you send it, that expression must vanish for every drive vector, which forces $S^H S = \mathbf{I}$. In other words a lossless network has a unitary scattering matrix: it rotates the vector of wave amplitudes without changing its length. The columns are then orthonormal, which is the algebraic way of saying that power sent into one port has to emerge from some port.

Allow loss and the equality relaxes to an inequality, $\mathbf{I} - S^H S \geq 0$, equivalent to requiring that every singular value of S be at most 1. A passive network is therefore a contraction on wave space — it can shrink the wave vector but never stretch it — and losslessness is the boundary case in which nothing shrinks at all. This singular-value form is the one to use in practice, because it remains the correct test when the network is lossy and S is no longer a normal matrix.

4.4 Eigenvectors of S: the drive patterns that keep their shape

Since a unitary S is diagonalisable as $S = U \Lambda U^H$ with every $|\lambda_k| = 1$, each eigenvector is a particular pattern of drives across the ports which emerges with the same shape it went in with, altered only by a phase $e^{j\theta}$. These patterns are a property of the matrix at whichever frequency you evaluated it, and like everything else in this section they are recomputed at each frequency of a sweep. It bears repeating that the unit circle appearing here lives in the plane of the eigenvalues of S and has nothing to do with the stability circle of the companion article. Once there is loss, the eigenvectors need not stay orthogonal, so passivity should be judged by singular values rather than eigenvalues, though the eigenvalues do still satisfy $|\lambda_k| \leq 1$.

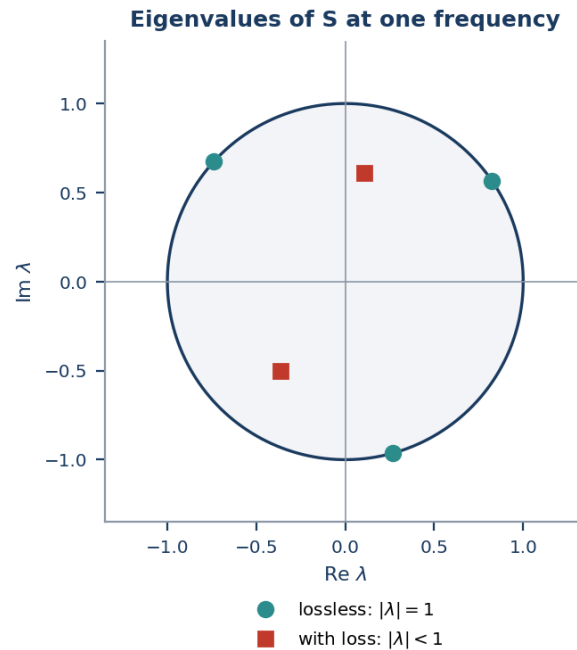


Figure 3 — Eigenvalues of S evaluated at one frequency. Lossless modes sit on the unit circle; loss pulls them inside. This is not a pole map, and the circle is not a stability boundary.

4.5 A worked feel for it: $25\ \Omega$ in a $50\ \Omega$ world

Solder a $25\ \Omega$ resistor in series between two $50\ \Omega$ ports and the scattering matrix comes out as $\begin{bmatrix} 0.2 & 0.8 \\ 0.8 & 0.2 \end{bmatrix}$, whose eigenvalues are $+1.0$ and -0.6 . Both numbers have a physical reading. Drive the two sides identically and both ends of the resistor sit at the same potential, so no current flows through it and nothing is dissipated: that even-mode pattern passes through untouched, which is what an eigenvalue of magnitude 1 means. Drive them in antiphase and the full difference appears across the resistor, which warms up, so that odd-mode pattern comes out smaller by a factor of 0.6. That is all an eigenvector of S amounts to — the drive pattern that keeps its shape — and the modulus of its eigenvalue tells you how much of it survives.

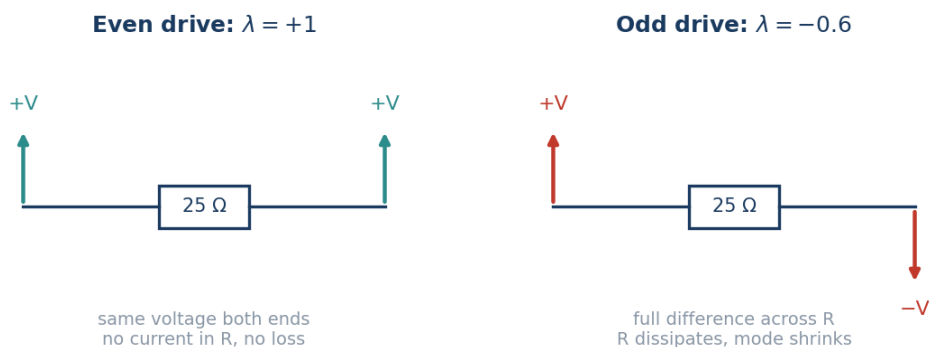


Figure 3b — The two eigenmodes of the $25\ \Omega$ example. The even drive puts no voltage across the resistor and is lossless; the odd drive puts all of it across the resistor and is attenuated.

5 Frequency response, impulse response and the shape of the curve

Everything so far has concerned the structure of the matrix. Plotting one of its entries against frequency introduces no new concept — $S_{21}(f)$ is just S_{21} evaluated at each f — but it does raise a new question, which is why the resulting curve has the shape it has. Answering that means looking inside the box, and the natural language inside is time rather than frequency.

5.1 Poles and zeros are eigenvalues of the internal state

For a network built from resistors, inductors and capacitors, every entry of the scattering matrix is a ratio of polynomials in s , say $S_{21}(s) = N(s)/D(s)$. The roots of the denominator are the poles and the roots of the numerator are the zeros. Those roots are not free-floating algebraic curiosities: collect every capacitor voltage and every inductor current into a state vector $\mathbf{q}$, write the network's dynamics as $d\mathbf{q}/dt = \mathbf{A}_c \mathbf{q} + \mathbf{b}\mathbf{x}$, and for a minimal model — one with no hidden states that cancel out of the output — the poles are exactly the eigenvalues of $\mathbf{A}_c$. A network that is both stable and passive has all of them strictly to the left of the $j\omega$ -axis, an axis which is simultaneously the fence they must stay behind and the road along which you measure.

5.2 Why a pole near the axis makes a peak

Walk up the $j\omega$ -axis watching $|S_{21}|$. A pole pair at $-\sigma \pm j\omega_0$ with a small σ means that when the measurement frequency reaches ω_0 the denominator nearly vanishes, so the response balloons. That is resonance, and the closer the pole sits to the axis the taller and narrower the peak becomes, with a quality factor $Q \approx \omega_0/2\sigma$. The mechanical analogy is exact enough to be worth keeping: small pushes delivered at the right rhythm build a large swing, and grabbing the ropes — adding loss, moving the pole away from the axis — kills it quickly. A zero sitting on the axis at $j\omega_z$ does the opposite, since the numerator vanishes exactly there and the response nulls out. Notch filters put zeros on the axis deliberately.

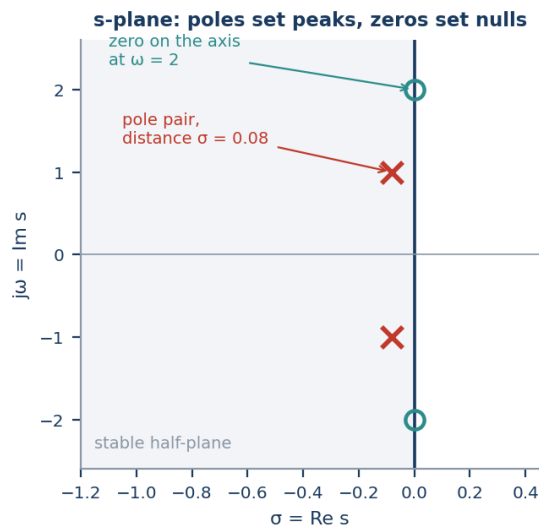


Figure 4 — An s -plane map of $H(s) = (s^2 + \omega_z^2)/(s^2 + 2\sigma s + \omega_0^2)$, with a pole pair close to the $j\omega$ -axis and a zero pair sitting on it.

The same example in numbers. Take $\sigma = 0.08$, $\omega_0 = 1$ and $\omega_z = 2$, so the poles are at $-0.08 \pm j0.997$ and the zeros at $\pm j2$. The quality factor is $Q \approx 1/(2 \times 0.08) \approx 6.3$. At $\omega = 1$ the numerator contributes 3 while the denominator collapses to $j0.16$, giving $|H| = 3/0.16 \approx 18.75$, a peak of roughly nineteen times. At $\omega = 2$ the numerator is exactly zero and so is the response. Move the pole out to $\sigma = 0.3$ and Q falls to about 1.7 while the peak drops to about 5: the same resonance, broader and shorter, because the pole has retreated from the axis.

5.3 The same information, seen in time

The impulse response of that network is a decaying oscillation, $e^{-\sigma t} \cos \omega_0 t$, and the two plots in Figure 5 contain identical information. The pole's distance from the axis, σ , sets the decay rate in the time-domain picture and the width of the peak in the frequency-domain picture; its height above the axis, ω_0 , sets the ringing frequency in one and the position of the peak in the other. A network that rings for a long time is a network with a tall narrow peak, and these are two descriptions of one behaviour rather than two behaviours. This is the pairing that deserves to be called two views of the network.

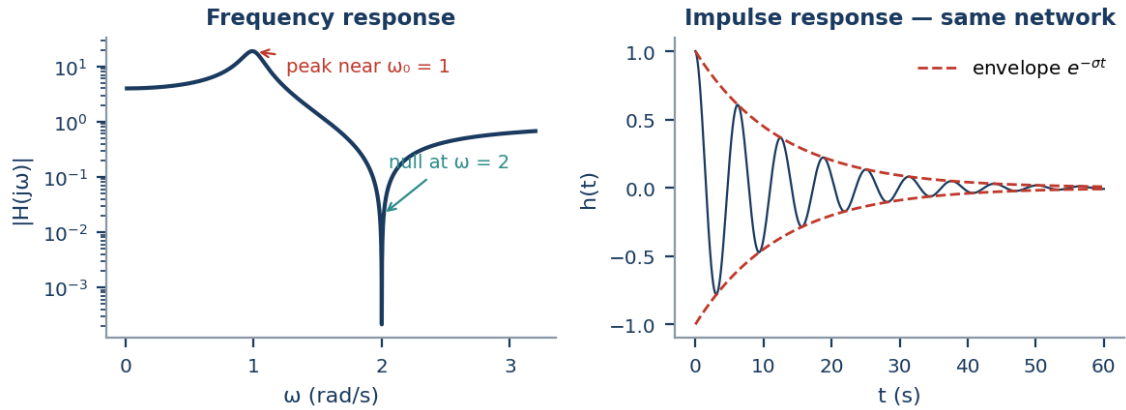


Figure 5 — One network, two plots. The frequency response peaks near ω_0 and nulls at ω_z ; the impulse response rings at ω_0 and decays at a rate set by the same σ that sets the width of the peak.

5.4 The digital twin

The companion article tells this story in discrete time, where a pole's radius in the z -plane plays the role that a pole's distance from the $j\omega$ -axis plays here. A pole close to the unit circle gives a mode λ^n that decays slowly and a correspondingly sharp peak in the frequency response, exactly as a pole close to the $j\omega$ -axis gives a slowly decaying $e^{-\sigma t}$ and a sharp peak here. Long ringing and narrow peaks go together in both worlds. What does not carry across is the eigenvalue picture of §4.4: the eigenvalues of S at a fixed frequency and the poles of a transfer function are eigenvalues of different matrices answering different questions, and the fact that both are sometimes compared against a unit circle is a coincidence of geometry rather than a connection of meaning.

6 From measurement to model: causality and passivity as constraints

What the laboratory actually hands you is a set of samples of $S(j\omega)$ taken at discrete frequencies, band-limited and noisy. Turning that into something a circuit simulator can use in the time domain means fitting a rational model, and the fit has to respect two physical constraints that the raw data only approximately encodes.

6.1 Causality and the Kramers-Kronig relations

A causal impulse response, one for which $h(t) = 0$ for $t < 0$, corresponds to a transfer function analytic in the right half-plane, and Cauchy's theorem then ties the real and imaginary parts of $H(j\omega)$ together as a Hilbert transform pair. These are the Kramers-Kronig relations, known in the network literature as Bode's attenuation-phase relations, and their practical content is that the magnitude and phase of a measured response are not independent quantities you may fit separately. One caveat deserves emphasis, since it is often stated loosely: the Kramers-Kronig relation itself requires only causality. The additional mirror rule that the real part is even and the imaginary part odd requires the impulse response to be real as well, which is true of VNA data taken on real voltages but false in general for complex-envelope models. The reliable way to enforce causality is structural rather than integral — fit a strictly proper rational function with all poles in the open left half-plane, as vector fitting does, and causality follows by construction.

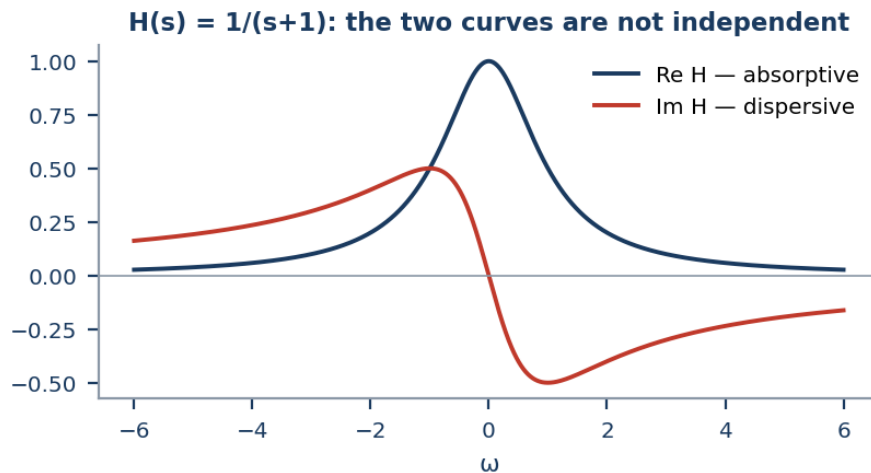


Figure 6 — The simplest causal pair, $H(s) = 1/(s+1)$. The absorptive and dispersive parts are locked to one another; you cannot adjust one without the other.

6.2 Passivity is the hard part

Causality restricts where the poles may sit, which is a condition on a finite set of numbers and is therefore easy to impose. Passivity is a condition at every frequency: it demands $S^H(j\omega)S(j\omega) \leq I$, or equivalently that the largest singular value never exceeds 1, and it must hold not merely at the frequencies you happened to sample but everywhere in between. An unconstrained fit will cheerfully sail past 1 in the gaps, and a model that does so manufactures energy and will make a transient simulation diverge, often long after the point where anyone was still suspicious of the model.

The standard remedy, due to Grivet-Talocia and Gustavsen, has two steps. The first is to locate every violation exactly rather than by sampling: the frequencies at which a singular value crosses 1 turn out to be the purely imaginary eigenvalues of a purpose-built $2N \times 2N$ Hamiltonian matrix, so a single eigenvalue computation finds all of them at once. This is a third and quite separate eigenvalue problem, distinct from the eigenvalues of S in §4 and from the poles in §5, and it is purely diagnostic. The second step is to perturb the residues by the smallest amount that pushes the maximum singular value back below 1, holding the poles fixed so that causality is not disturbed, which is a constrained quadratic program. Only a model that is both causal and passive is safe to hand to a time-domain simulator.

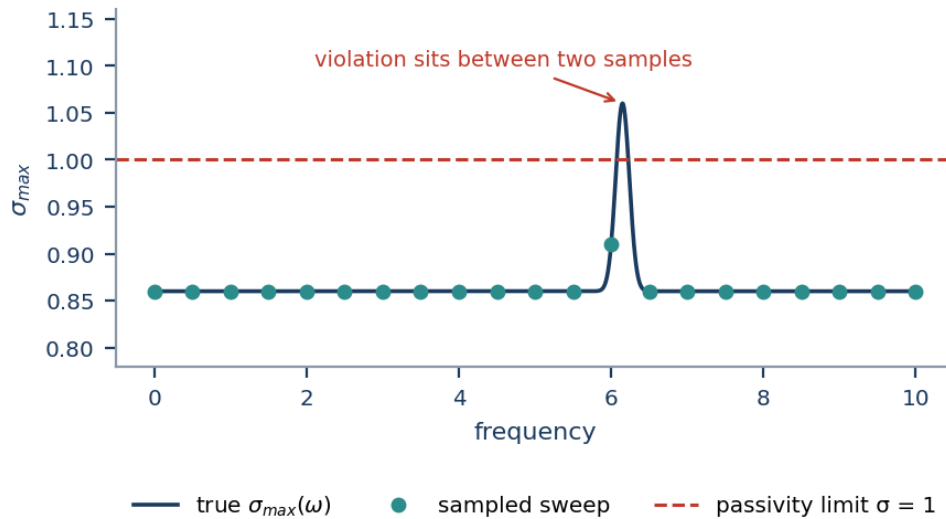


Figure 7 — A violation hiding between samples. Every sampled point passes the test while the true maximum singular value exceeds 1 in the gap; only the all-frequency Hamiltonian test catches this.

7 The linear algebra underneath

Four standard results carry most of the weight in this report, and it is worth seeing them collected in one place with the use each is put to.

Theorem	What it says	Where it is used here
Spectral theorem for Hermitian matrices	Real eigenvalues and an orthonormal eigenvector basis	Passivity via Z_H stays a real eigenvalue problem, so pass or fail is decided by the sign of a real number
Unitary diagonalisation	$S = U\Lambda U^H$ with every $ \lambda = 1$	The eigenvector picture of a lossless network in §4.4 and Figure 3
Normal matrices	$AA^H = A^HA$ implies unitary diagonalisability	A reciprocal lossless S has orthogonal eigenmodes; a lossy one need not, which is why singular values are the safer test
Positive semidefiniteness	All eigenvalues ≥ 0 is equivalent to $v^H M v \geq 0$ for every v	Passivity written as a quadratic form in the drive vector
General spectra	No constraint on modulus	The poles in A_C may sit anywhere in the left half-plane, and their distance from the axis sets Q

Read on the bench, these say that a passivity check reduces to asking whether a real number is non-negative; that a lossless network conserves the energy in each column of S , so a column energy above 1 is unphysical; and that lossless reciprocal modes cooperate while lossy ones can interfere, which is the practical reason to reach for singular values. Horn and Johnson, Strang, and Shankar all give the proofs.

8 Converting between Z, Y and S, and the Smith chart

With a single real reference impedance common to all ports, the conversions are compact. The admittance matrix is the inverse of the impedance matrix, $Y = Z^{-1}$. The scattering matrix follows from the impedance matrix as $S = (Z + Z_0 I)^{-1}(Z - Z_0 I)$, and the relation inverts to $Z = Z_0(I + S)(I - S)^{-1}$. This is a matrix Möbius transformation, and it carries the half-plane $\text{Re}(Z) \geq 0$ onto the unit ball $|S| \leq 1$: passive networks become contractions, and purely reactive ones land on the boundary. Different real reference impedances at different ports are handled by inserting the appropriate per-port $\sqrt{Z_0}$ normalisation factors, and nothing conceptual changes. A complex reference impedance is a different matter, and §8.1 sets out why.

8.1 Three conventions that agree except at a complex reference

The literature carries three definitions of the incident and reflected amplitudes, and they agree everywhere except when the reference impedance is complex. Travelling voltage waves, used in §4.2 and throughout signal-integrity practice, keep the same reference in both expressions, $V^\pm = (V \pm Z_0 I)/2$, and give a reflection coefficient $\Gamma = (Z - Z_0)/(Z + Z_0)$. Kurokawa's power waves conjugate the reference in the outgoing term, $b = (V - Z_0^* I) / 2\sqrt{\text{Re } Z_0}$, and give $\Gamma = (Z - Z_0^*)/(Z + Z_0)$. Marks and Williams' pseudo-waves are the careful generalisation for a reference impedance that is itself complex and frequency-dependent, as on a lossy line or on wafer.

With a real Z_0 the conjugate does nothing and all three coincide, which is why the distinction can be set aside for most 50 Ω coaxial and PCB work. With a complex reference they genuinely differ, and the difference is physical rather than cosmetic: a power-wave Γ of zero means the port is conjugate matched, so power transfer is maximised, whereas a travelling-wave Γ of zero means the port is reflectionless, so no backward wave exists on the line at all. Those are two different terminations, and the conjugate is not a normalisation factor you can absorb. State which convention you are using whenever Z_0 is not real.

8.2 The Smith chart is this transformation drawn for one port

For a single port the matrices become scalars and the map reduces to $\Gamma = (z - 1)/(z + 1)$ with $z = Z/Z_0$ the normalised impedance. That is the Smith chart, and its familiar grid of circles is nothing more than the image of the constant-resistance and constant-reactance lines under this bilinear map. The rim $|\Gamma| = 1$ is the one-port case of the circle in Figure 3, carrying the purely reactive terminations, with everything passive inside it. A short circuit sits at $\Gamma = -1$, a matched load at the centre, an open at $+1$; 25 Ω normalises to $z = 0.5$ and lands at $\Gamma = -1/3$, a VSWR of 2, while 100 Ω lands symmetrically at $+1/3$. Left of centre is low impedance, right of centre is high, and the rim is reactive. There is no useful N-port generalisation of the picture, which is why multiport work uses the matrix formulae above instead.

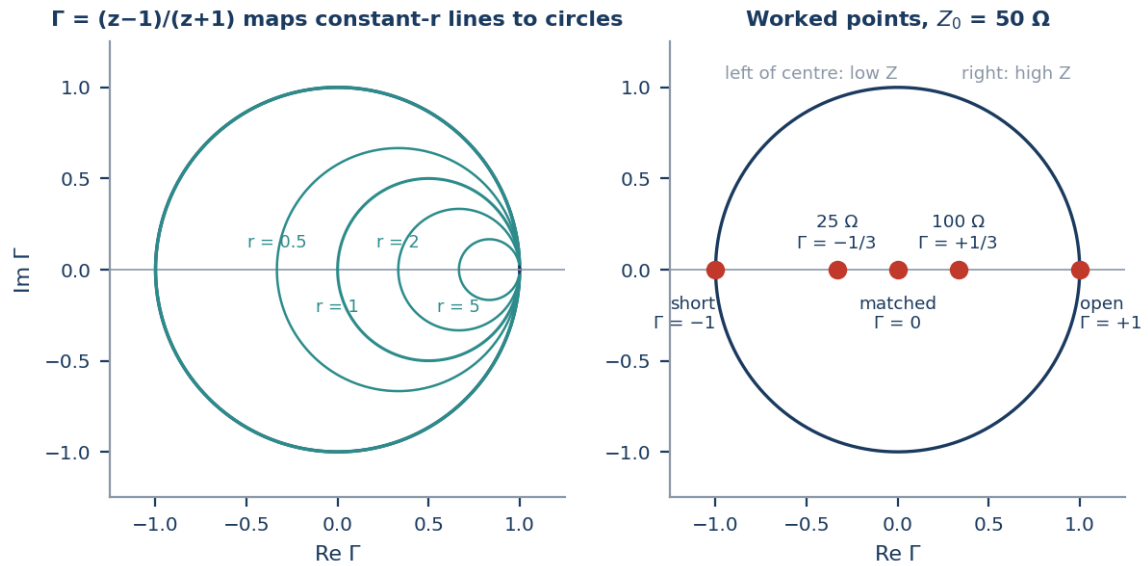


Figure 8 — The bilinear map behind the Smith chart. On the left, constant-resistance lines becoming nested circles tangent at the open-circuit point; on the right, the worked terminations for $Z_0 = 50 \Omega$.

Physical property	In Z or Y form	In S form
Reciprocal	$Z = Z^T$	$S = S^T$, a symmetric matrix
Passive	$Z_H \geq 0$	$I - S^H S \geq 0$, all singular values ≤ 1
Lossless	$Z_H = 0$	$S^H S = I$, unitary, all $ \lambda = 1$
Y exists at all	$\det(Z) \neq 0$	$\det(I - S) \neq 0$

Table 1 — Physical property against matrix property, assuming one real reference impedance common to every port.

9 Mixed-mode S-parameters: differential and common in one matrix

9.1 It is a change of basis, not a new measurement

A differential pair presented to a vector network analyser is a four-port: two single-ended ports at the near end, one on each conductor of the pair, and two more at the far end. The 4×4 matrix that comes back says nothing directly about differential behaviour, because the differential and common modes are combinations of the single-ended variables rather than variables in their own right. Defining them at each pair as $v_d = v_p - v_n$ and $v_c = (v_p + v_n)/2$, with the matching current definitions $i_d = (i_p - i_n)/2$ and $i_c = i_p + i_n$ that keep the power bookkeeping consistent, the change of variables is linear and can be written as a single matrix.

With the single-ended ports numbered so that 1 and 2 are the two conductors at the near end and 3 and 4 the two at the far end, and the mixed-mode vector ordered as (d_1, d_2, c_1, c_2) , that matrix is $M = (1/\sqrt{2})[[1, -1, 0, 0], [0, 0, 1, -1], [1, 1, 0, 0], [0, 0, 1, 1]]$ and the mixed-mode scattering matrix is $S_{mm} = M S M^T$. The $1/\sqrt{2}$ makes M orthogonal, so $M^{-1} = M^T$ and the operation is a similarity transform — the same construction as §2, one linear operator viewed in a different basis. Nothing is measured twice: mixed-mode parameters are computed from the four-port data and are exactly as trustworthy as it is.

Why the basis change costs nothing. Because M is orthogonal, it is a rotation of the wave space, and rotations preserve eigenvalues and singular values. Every test in §3, §4 and §6 — reciprocity, passivity, losslessness, the Hamiltonian search — returns the same answer whether you run it on the single-ended matrix or the mixed-mode one. The transformation reorganises the information into the combinations an engineer actually cares about; it neither adds any nor destroys any.

9.2 The four quadrants and what each one is for

Partitioned into 2×2 blocks, the mixed-mode matrix reads as four separate stories.

Block	What it describes	Why you look at it
S_{dd}	Differential in, differential out	The actual signal path. S_{dd21} is the insertion loss that sets how much equalisation a link needs; S_{dd11} is the differential return loss
S_{cc}	Common in, common out	How common-mode energy propagates. Governs whether a common-mode choke or a ferrite will achieve anything, and how much common-mode current reaches a cable
S_{cd}	Differential in, common out	Your own signal leaking into common mode. Common-mode current on an attached cable radiates efficiently, so this is the quadrant that fails an EMC scan
S_{dc}	Common in, differential out	External common-mode noise arriving as differential noise at the receiver. It lands directly on the eye and eats margin

The structural point is the one worth remembering. If the two halves of the pair are genuinely identical, the off-diagonal blocks vanish and S_{mm} is block diagonal: differential and common mode propagate independently and never exchange energy. Every non-zero entry in S_{cd} or S_{dc} is therefore a calibrated measure of how asymmetric the pair really is, which is a far more useful thing to hand a layout engineer than a vague complaint about EMC.

The four quadrants of a mixed-mode S-matrix

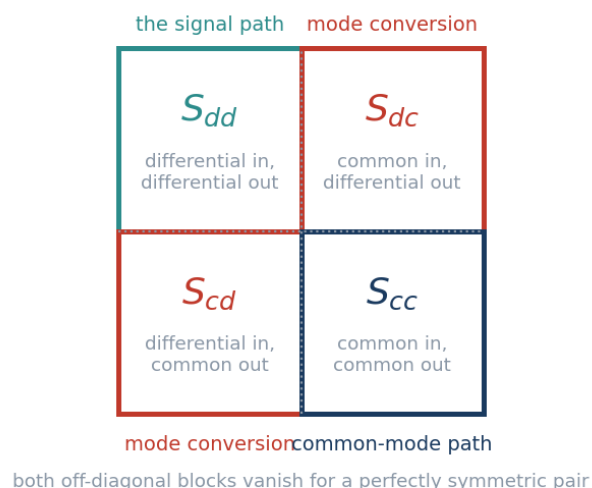


Figure 9 — The four quadrants. The diagonal blocks are the differential and common channels; the off-diagonal blocks exist only to the extent that the pair is asymmetric.

9.3 What breaks the symmetry, and by how much

The cleanest case to work through is intra-pair skew. Suppose the two conductors are identical in every respect except that one is electrically longer by Δt . Splitting the delay symmetrically between them and driving the pair differentially, the far-end voltages are $H e^{-j\omega\Delta t/2}$ and $-H e^{+j\omega\Delta t/2}$, so the differential and common outputs are

$$|S_{dd21}| = |H| \cos(\pi f \Delta t), \quad |S_{cd21}| = |H| \sin(\pi f \Delta t) \quad (15)$$

— the energy the differential mode loses turns up in the common mode, exactly and without waste. Put numbers on it and the asymmetry between the two effects is striking. At 14 GHz, 2 ps of skew converts –21 dB of the through signal into common mode while costing the differential path only 0.03 dB. At 26.6 GHz, the Nyquist frequency of a 53 GBd lane, 5 ps of skew converts –7.8 dB — nearly a fifth of the wave amplitude — for a differential penalty of 0.8 dB.

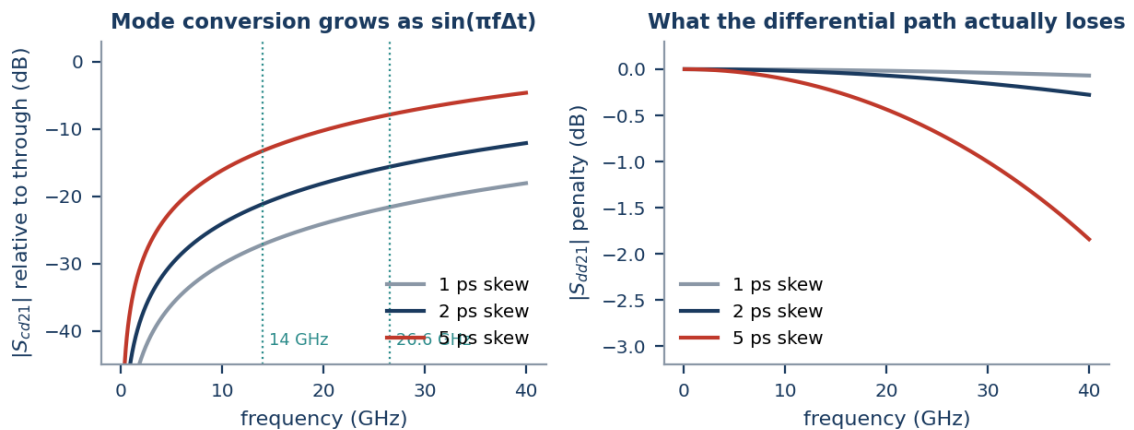


Figure 10 — Intra-pair skew, seen two ways. Mode conversion (left) rises steeply with frequency and skew, while the differential insertion loss (right) barely moves. You cannot detect skew by looking at S_{dd21} .

The practical reading. S_{dd21} is almost blind to skew, so a channel can look perfectly healthy on the plot everyone shows in review while radiating badly. S_{cd21} is the sensitive instrument, and it is the one to put a limit line on. The usual causes are unequal path lengths where a pair turns a corner or dodges a via, asymmetric via barrels, antipads and connector footprints, and — most awkwardly — fibre weave, where one conductor of the pair happens to run over glass bundles and the other over resin, giving the two a different effective permittivity. Weave-induced skew accumulates with length and is why long pairs are routed at a small angle to the weave, or on spread-glass laminates.

One reassurance is worth stating explicitly, because the word 'conversion' sounds alarming. Mode conversion is not loss. Energy moves from one mode to the other and the four-port remains perfectly passive; the singular values of S are untouched, as §9.1 guarantees. The trouble with converted energy is not that it has been destroyed but that it is now in a mode the receiver does not read and an enclosure does not contain.

9.4 The port-numbering trap

A Touchstone file records the numbers but not what the ports were physically connected to, and two conventions are in common use: ports 1 and 2 as the near-end pair with 3 and 4 as the far-end pair, or ports 1 and 3 as the two ends of one conductor with 2 and 4 as the other. Applying the transformation for one convention to data taken in the other permutes the blocks — most visibly, it swaps S_{dd} with S_{cc} — and produces plots that look like a catastrophically bad channel. The sanity check takes a few seconds: S_{dd21} should be the least lossy path on the chart, S_{cc21} should be worse, S_{dd11} should be well behaved, and both conversion terms should be small. If S_{dd21} looks like rubbish while S_{cc21} looks pristine, you have the permutation the wrong way round rather than a broken board.

10 Where this shows up in practice

In RF work the two numbers most often quoted are $|S_{11}|$, which expresses mismatch and return loss, and $|S_{21}|$, which expresses insertion loss or gain. A vector network analyser produces S directly, and models travel between tools as Touchstone .s2p or .snp files. The singular-value bound is the gate that electromagnetic extractions have to pass before they are trusted, and de-embedding a fixture is done by cascading in the ABCD domain and converting back.

In high-speed digital work the same objects appear under a different name. PCB traces, connectors and packages are characterised as multiport S — very often as a four-port differential structure — and enter SPICE or IBIS-AMI flows as macromodels used to predict eye diagrams and bit error rates. Those macromodels come from the vector fitting of §6, and the passivity enforcement described there is what makes them safe to simulate, with the mixed-mode decomposition of §9 turning the four-port into the differential and common channels a link budget is actually written in. And the equaliser taps in a SerDes, whether feed-forward or decision-feedback, are the solution of exactly the Hermitian positive-semidefinite least-squares problem that the companion article treats in its section on optimal filter taps.

11 Summary and a checklist for laboratory data

Claim about the network	How you test it	Which boundary it involves
Reciprocity	$Z = Z^T$, or $S = S^T$	Symmetry of the matrix; no boundary
Passivity	$Z_H \geq 0$, equivalently $\sigma(S) \leq 1$ at every frequency	Positive semidefiniteness and contraction
Losslessness	$S^H S = I$	The unit circle in the plane of the eigenvalues of S
A resonance or a null	A pole close to the $j\omega$ -axis, or a zero on it; $Q \approx \omega_0/2\sigma$	Distance from the $j\omega$ -axis in the s -plane
Fit is safe to simulate	Left-half-plane poles plus Hamiltonian passivity enforcement	No right-half-plane poles, $\sigma \leq 1$ everywhere
Pair is symmetric	S_{cd21} small across the band	Off-diagonal mixed-mode blocks vanish for a symmetric pair

Before you trust an S-parameter model. First, check reciprocity: unless the device contains ferrite or active circuitry, S should equal its own transpose, and the size of the antisymmetric part is a direct measure of measurement error. Second, check causality: the fitted poles should all lie in the left half-plane, and a Kramers-Kronig spot check on magnitude against phase will catch gross violations. Third, check passivity properly, using the all-frequency Hamiltonian test rather than a sweep through the sample points, since that is precisely where the violations hide. Fourth, sanity-check the resonances: every peak should correspond to a left-half-plane pole pair whose Q matches the observed width, and every null to a nearby zero. Fifth, check that the reference impedance is consistent — one real Z_0 , usually $50\ \Omega$ — before applying any of the conversion formulae in §8.

Glossary

Causal. A response for which $h(t) = 0$ when $t < 0$. Equivalent to analyticity in the right half-plane, and the source of the Kramers-Kronig link between real and imaginary parts.

Contraction. A matrix with $M^H M \leq I$, so that no singular value exceeds 1. A lossy passive S is a contraction and a lossless one sits on its boundary.

Eigenvector of S . A drive pattern across the ports that emerges with its shape unchanged at the frequency where S was evaluated. Not a pole.

Hamiltonian passivity test. A $2N \times 2N$ construction whose purely imaginary eigenvalues mark the frequencies at which a singular value of S crosses 1.

Kramers-Kronig relations. The Hilbert-transform pairing of real and imaginary parts forced by causality; Bode's attenuation-phase relations.

Möbius map. A transformation of the form $(ax + b)/(cx + d)$. Here it carries $\text{Re}(Z) \geq 0$ onto $|S| \leq 1$.

Pole and zero. Roots of the denominator and numerator of $S_{21}(s)$. A pole near the $j\omega$ -axis produces a peak; a zero on it produces a null.

Travelling (voltage) wave. The forward and backward voltage on a line, $V^\pm = (V \pm Z_0 I)/2$, in volts. $S_{ij} = V_i^- / V_j^+$.

Power wave. Kurokawa's normalisation, $a = (V + Z_0 I)/2\sqrt{\text{Re } Z_0}$, $b = (V - Z_0^* I)/2\sqrt{\text{Re } Z_0}$, chosen so that $|a|^2$ and $|b|^2$ are powers even for a complex reference. Note the conjugate in b : with Z_0 real it does nothing, and power waves are then the travelling waves divided by $\sqrt{Z_0}$.

Pseudo-wave. The Marks and Williams generalisation used when the reference impedance is itself complex and frequency-dependent.

Mixed-mode S-parameters. $S_{mm} = M S M^T$ with M orthogonal; the same four-port operator written in differential and common coordinates.

Mode conversion. The off-diagonal mixed-mode blocks S_{cd} and S_{dc} . Non-zero only in proportion to the pair's asymmetry, and not a loss mechanism — the energy changes mode rather than disappearing.

Q . Approximately $\omega_0/2\sigma$. A high Q means a pole close to the $j\omega$ -axis and a response that rings for a long time.

State matrix A_c . The matrix in $d\mathbf{q}/dt = A_c \mathbf{q} + \mathbf{b}\mathbf{x}$; for a minimal model its eigenvalues are the poles.

Vector fitting. A rational pole-residue fit to sampled $S(j\omega)$, used to build models for time-domain simulation.

References

- [1] Pozar, Microwave Engineering, 4th ed., Wiley, 2011.
- [2] Kurokawa, 'Power waves and the scattering matrix', IEEE Trans. MTT, 1965.
- [3] Bogatin, 'How not to be confused by S-parameters', Signal Integrity Journal, 2020 — the signal-integrity view, in voltage waves throughout.

- [4] Marks and Williams, 'A general waveguide circuit theory', J. Res. NIST, vol. 97, no. 5, 1992 — pseudo-waves for a complex, frequency-dependent reference.
- [5] Horn and Johnson, Matrix Analysis, 2nd ed., Cambridge, 2012.
- [6] Strang, Linear Algebra and Its Applications, 4th ed., Cengage, 2006.
- [7] Shankar, Basic Training in Mathematics, Springer, 1995.
- [8] Bode, Network Analysis and Feedback Amplifier Design, Van Nostrand, 1945.
- [9] Van Valkenburg, Introduction to Modern Network Synthesis, Wiley, 1960.
- [10] Gustavsen and Semlyen, 'Rational approximation of frequency domain responses by vector fitting', IEEE Trans. Power Delivery, 1999.
- [11] Grivet-Talocia and Gustavsen, Passive Macromodeling, Wiley, 2016.
- [12] Rytting, 'Network analyzer error models and calibration methods', Agilent/HP application note — background on SOLT and TRL.
- [13] Niknejad, EECS 242 lecture notes, UC Berkeley.