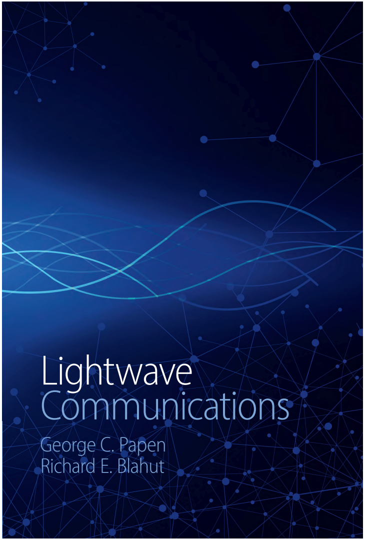
The Lightwave Channel: Particles or Waves

ISTC 2018

Richard E. Blahut

Richard Blahut is Emeritus with the Department of Electrical and Computer Engineering, University of Illinois, and is now with the Department of Electrical and Systems Engineering, University of Pennsylvania

In collaboration with George Papen


Lightwave
Communications
George C. Papen
Richard E. Blahut

Background

- Within information theory, wave and particle channels seem to be different
- Within physics, electromagnetic waves are granular at low signal levels
- Quantum theory unifies waves and particles

Background

- Information theory is a subject that exists separate from physics
- Within information theory, the Poisson transform unifies wave channels and particle channels
- Within physics, this unification would be called a semiclassical analysis

Observation

- Shannon bandlimited capacity

$$C = B \log(1 + S/N)$$

for maximum-entropy, worst-case noise.

- Energy per bit

$$E_b/N_0 \geq -1.6dB$$

- These statements are mathematically complete.
- Or are they?

Observation

- Shannon bandlimited capacity

$$C = B \log(1 + S/N)$$

- Energy per bit

$$E_b/N_0 \geq -1.6dB$$

- These statements are mathematically complete.
- These statements are not physically complete.
- These statements fail to account for granularity.

Granularity

- At 2×10^{10} Hertz (20 Gigahertz)
  One Watt equals 7.5×10^{22} photons per second
- At 2×10^{14} Hertz (1.5 microns)
  One nanowatt equals 7.5 photons per nanosecond

Information-Theoretic Channel Models

- Waveform Channel
  - Shannon (1948)
- Particle Channel
  - Many authors
- Wave/Particle Channel
  - Gordon-(Forney) Conjecture (1964)
  - Gordon formula (1962)
  - Blahut-Papen (2018)
- Quantum Channel
  - Von Nuemann (1932)
  - Holevo Bound (1972)
  - Phase Sensitive Channel — many authors (2014)

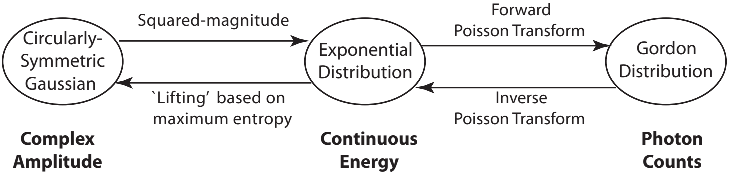
The Physics of Photons and Waves

- "Particle" and "complex baseband signal" have physical meaning using instead the terms "photon" and "wave. "
  - Photons are each associated with a constant E called the *energy*.
  - Waves are associated with a constant f called the *carrier frequency*.
- These constants are related by $E = hf$ where the scaling term h is called Planck's constant.

Objective

- Relate the channel capacity of a discrete “particle” channel to the channel capacity of a continuous “complex-baseband” channel.
- Provide a common framework to give the capacity of Poisson channel in a small-signal regime and the Shannon capacity in a large signal regime.
- Shannon capacity should be a large-signal emergent fluid model from the capacity of a particle stream.

Approach - Relate Continuous and Discrete Maximum-Entropy Distributions



Maximum Entropy Distribution for Particles

- The probability constraints are that $\sum_{\mathrm{m}=0}^{\infty} p_{\underline{\mathrm{m}}}(\mathrm{m}) = 1$, and the p_m are nonnegative.
- The maximization is constrained by a finite mean $E[m] = \mathrm{M}$.

Theorem

The maximum-entropy probability mass function is

$$p_{\underline{\mathrm{m}}}(\mathrm{m}) = \frac{1}{1+\mathrm{M}}\left(\frac{\mathrm{M}}{1+\mathrm{M}}\right)^{\mathrm{m}} \qquad \mathrm{m} = 0, 1, \ldots.$$

Proof.

This is a standard maximization using Lagrange multipliers. □

Note: This is not a Poisson distribution.

The Gordon Distribution

- This probability mass function is a geometric probability mass function called the *Gordon distribution*
- The entropy of the Gordon distribution is

$$H = \log_e(1 + \mathrm{M}) + \mathrm{M}\log_{\mathrm{e}}(1 + 1/\mathrm{M})$$

- The two terms suggest the wave and particle nature of a signal.
  - When M is large $H \approx \log_e(1 + \mathrm{M})$. This is the entropy of a continuous exponential probability density function with mean M.
  - When M is is small $H \approx M - M\log_e \mathrm{M}$. This is the small-signal expansion of the entropy of a Poisson probability mass function with mean M.
- Two limiting forms of the Gordon distribution reveal the particle and wave properties.

Special Distributions

- Probability density function
  - Maximum entropy — Real or complex gaussian
  - Convolution invariance — Real or complex gaussian
- Probability mass function
  - Maximum entropy — Gordon distribution
  - Convolution invariance — Poisson distribution

The Poisson Transform

- Let $f_{\underline{\mathsf{E}}}(\mathsf{E})$ be any continuous probability density function on the nonnegative reals. Then

$$p_{\underline{\mathsf{m}}}(\mathsf{m}) = \int_0^\infty \frac{\mathsf{E}^\mathsf{m}}{\mathsf{m}!} e^{-\mathsf{E}} f_{\underline{\mathsf{E}}}(\mathsf{E}) \mathrm{d}\mathsf{E},$$

  is a probability mass function on the nonnegative integers.
- The inverse Poisson transform maps probability mass functions to real probability density functions.

Properties of the Poisson Transform

- The Poisson transform (and its inverse) reveal the parallel roles for the wave model and the particle model of a signal.
- By analogy with the symbolic expression for the Fourier transform

$$s(t) \quad \longleftrightarrow \quad S(f)$$

  the Poisson transform is expressed symbolically as

$$f(x) \quad \leftrightsquigarrow \quad p(m).$$

Examples

- The Poisson distribution is the Poisson transform of a Dirac delta function.

$$\delta(E - M) \quad \leftsquigarrow\!\!\rightsquigarrow \quad \frac{\mathsf{M}^{\mathsf{m}}}{\mathsf{m}!} e^{-\mathsf{M}}.$$

- The Gordon distribution is the Poisson transform of a maximum-entropy exponential probability density function.

$$e^{-E/M} \quad \leftsquigarrow\!\!\rightsquigarrow \quad \frac{1}{1+\mathsf{M}} \left(\frac{\mathsf{M}}{1+\mathsf{M}} \right)^{\mathsf{m}}$$

Composite Distribution formed by the Poisson Transform

- When viewed as a Poisson transform, the Gordon distribution is the composite of:
  - The uncertainty caused by the random arrival times of particles,
  - Maximum statistical uncertainty expressed by an exponential distribution.
- The composite effect determined by
  - Considering the effect of the channel (always present).
  - Overlying maximum-entropy statistical uncertainty using the Poisson transform to give the Gordon distribution.

Lifting the Energy to the Complex Amplitude

- The square of a constant complex amplitude over a finite-time interval T is the energy E.
- The square-root of the maximum-entropy exponential distribution is not a maximum-entropy distribution
- This failure is remedied by the assertion that the square-root of the energy is complex.
- Equivalently, the energy is the sum of two squared terms, not one.

Position and Momentum

- A particle is described by both a position and a momentum and has both potential energy and kinetic energy.
- An approach that does not explicitly account for these two degrees of freedom of the particle fails to generate a maximum-entropy distribution.

Inverting the Exponential Distribution

- $x + \mathrm{i}\mathrm{y}$ is gaussian $\longrightarrow x^2 + y^2$ is exponential
- What is square root of exponential?
  - $\sqrt{E}$ is not maximum entropy
  - $\sqrt{E}$ does not convolve
- Energy must be expressed as the complex factorization

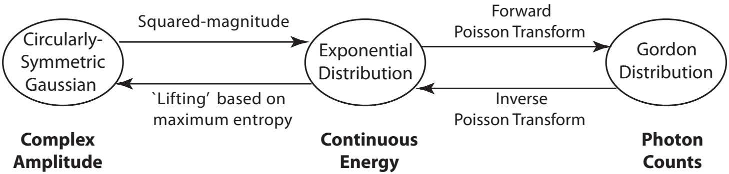
  $$E = (x + \mathrm{i}y)(x - \mathrm{i}y)$$

  or

  $$E = x^2 + y^2.$$

- This expression, hinted at by the form of the composite Gordon distribution, is an unavoidable consequence of respect for the maximum entropy principle.
- The maximum-entropy distribution of the complex amplitude $A = A_I + \mathrm{i}A_Q = |A|e^{\mathrm{i}\phi}$ is a circularly-symmetric gaussian.

Summary



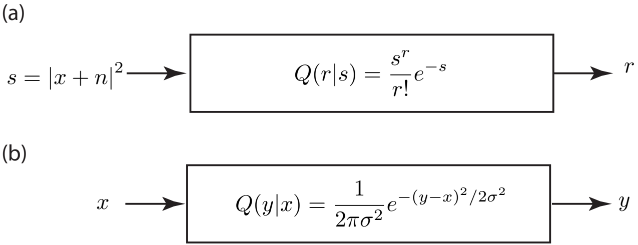
- A unified information-theoretic framework for both wave channels and particle channels
  - Consistent with quantum theory.
  - Semiclassical abridgement of quantum theory.

The Additive Noise Particle Channel

- A discrete memoryless channel is composed of contiguous equal time intervals each with a complex sinusoid
  - Average signal power constrained.
  - Independent, additive complex gaussian noise in each interval.
- The random variable $\underline{r}$ denotes the total number of received particles in an interval.
  - Poisson transform of wave intensity
- Considering only the discrete-particle aspect of the signal, the random received signal $\underline{r}$ is the sum of particle counts for the signal $\underline{s}$ and the independent additive noise $\underline{n}$

$$\underline{r} = \underline{s} + \underline{n}.$$

Two Channel Models



Mutual Information and Capacity

- Using $\underline{r} = \underline{s} + \underline{n}$, the mutual information $I(\underline{s}; \underline{r})$ is

$$I(\underline{s}; \underline{r}) = H(\underline{r}) - H(\underline{r}|\underline{s}) = H(\underline{s} + \underline{n}) - H(\underline{n}),$$

  where $H(\underline{r}|\underline{s})$ is equal to $H(\underline{n})$ because the signal $\underline{s}$ and the noise $\underline{n}$ are independent.
- The conditional probability $p(\underline{r}|\underline{s})$ that determines the entropy $H(\underline{r})$ is a conditional Poisson distribution

$$\mathrm{C} = \max I(\underline{s}; \underline{r}),$$

Maximizing the Mutual Information for the Particle Channel

- Distribution $p(\mathsf{r})$ maximizing entropy $H(\underline{\mathsf{r}})$ is the Gordon distribution given by
$$p(\mathsf{r}) = \frac{1}{1+\mathsf{Z}}\left(\frac{\mathsf{Z}}{1+\mathsf{Z}}\right)^{\mathsf{r}},$$
with entropy $H(\underline{\mathsf{r}}) = (1+\mathsf{Z})\log(1+\mathsf{Z}) - \mathsf{Z}\log\mathsf{Z}$ where Z is the expected number of received particles $\mathsf{E}+\mathsf{N}_0$.
- Capacity would be achieved for a prior $p(\mathsf{s})$ such that $p(\mathsf{s}+\mathsf{n})$ is a maximum-entropy Gordon distribution with $\mathsf{Z}=\mathsf{E}+\mathsf{N}_0$.

The Channel Capacity

- When $\underline{r}$ and $\underline{n}$ are each Poisson, and $\underline{r} = \underline{s} + \underline{n}$, then $\underline{s}$ must be Poisson as well.
- When the encoder generates an input distribution for the continuous signal energy that is the maximum-entropy exponential distribution, the corresponding discrete distribution for the number of transmitted particles is a Gordon distribution.
- Mean signal E is the difference between the mean number of received particles Z and the mean number of noise particles N_0 added by the channel.
- Using $\mathsf{Z} = \mathsf{E} + \mathsf{N}_0$, the maximum entropy $H(\underline{r})$ at the output of the channel is

$$H(\underline{r}) = g(\mathsf{Z}) \quad = \quad g(\mathsf{E} + \mathsf{N}_0).$$

The Channel Capacity

- Using the precceding expression and $H(\underline{\mathsf{n}})$, the capacity of the channel is

$$\mathsf{C} = H(\underline{\mathsf{r}}) - H(\underline{\mathsf{n}}) = g(\mathsf{E} + \mathsf{N}_0) - g(\mathsf{N}_0),$$

- This is the single-letter capacity C of the noisy time-discrete particle channel in units of bits/symbol.

$$\boxed{\mathsf{C} = \log_2\left(1 + \frac{\mathsf{E}}{1 + \mathsf{N}_0}\right) + (\mathsf{E} + \mathsf{N}_0)\log_2\left(1 + \frac{1}{\mathsf{E} + \mathsf{N}_0}\right) - \mathsf{N}_0 \log_2\left(1 + \frac{1}{\mathsf{N}_0}\right)}$$

- E is the mean number of signal counts
  - $E = \mathsf{E}hf$ is the mean signal energy
- N_0 is the mean number of noise counts
  - $N_0 = \mathsf{N}_0 hf$ is the mean noise energy

Gordon Noiseless Capacity Formula

The bandlimited capacity $\mathcal{C}$, in bits per second, of a *noiseless* particle channel with time-averaged particle arrival rate R, in bits per second,

$$\mathcal{C} \quad = \quad B \log_2 \left(1 + \frac{\mathsf{R}}{B}\right) + \mathsf{R} \log_2 \left(1 + \frac{B}{\mathsf{R}}\right)$$

with B in Hertz.

The Shannon Capacity

- The Shannon capacity based on waves can be derived directly from the particle capacity.
- When both E and N_0 are much larger than one, the entropy of a Gordon distribution $g(x)$ approaches the entropy of an exponential distribution.
- Replace mean particle counts E and N_0 by mean continuous energy E and N_0
- Replacing $g(\mathrm{m})$ in by $1 + \log x$ gives

$$\begin{aligned} \mathsf{C} &= H(\underline{\mathrm{r}}) - H(\underline{\mathrm{n}}) \\ &= 1 + \log(E + N_0) - \left(1 + \log N_0\right) \\ &= \log_2\left(1 + \frac{E}{N_0}\right) \quad \text{bits per symbol} \end{aligned}$$

which is the single-letter capacity based on waves.

Comments and Conclusion

- The Poisson transform (lifted) provides a unified framework for wave channels and particle channels
- The duality of channel capacity of particle and wave channels is one example

Information Theory Hierarchy

1. Classical information theory
   (Shannon, *et al*)

2. Semiclassical information theory
   (Gordon, Mandel, Wolf, Forney, *et al*)

3. Quantum information theory
   (von Neumann, Holevo, Shapiro, *et al*)