

PAPER • OPEN ACCESS

Light scattering as a Poisson process and first-passage probability

To cite this article: Claude Zeller and Robert Cordery *J. Stat. Mech.* (2020) 063404

View the [article online](#) for updates and enhancements.



IOP | ebooksTM

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

PAPER: Interdisciplinary statistical mechanics

Light scattering as a Poisson process and first-passage probability

Claude Zeller¹ and Robert Cordery^{2,3}

¹ Claude Zeller Consulting LLC, Oceanside, OR 97134, United States of America

² Fairfield University, 1073 North Benson Rd., Fairfield, CT 06824, United States of America

E-mail: czeller@ieee.org and rcordery@fairfield.edu

Received 25 June 2019

Accepted for publication 6 March 2020

Published 24 June 2020



Online at stacks.iop.org/JSTAT/2020/063404

<https://doi.org/10.1088/1742-5468/ab811f>

Abstract. A particle entering a scattering and absorbing medium executes a random walk through a sequence of scattering events. The particle ultimately either achieves first-passage, leaving the medium, or it is absorbed. The Kubelka–Munk model describes a flux of such particles moving perpendicular to the surface of a plane-parallel medium with a scattering rate and an absorption rate. The particle path alternates between the positive direction into the medium and the negative direction back towards the surface. Backscattering events from the positive to the negative direction occur at local maxima or peaks, while backscattering from the negative to the positive direction occur at local minima or valleys. The probability of a particle avoiding absorption as it follows its path decreases exponentially with the path-length λ . The reflectance of a semi-infinite slab is therefore the Laplace transform of the distribution of path-length that ends with a first-passage out of the medium. In the case of a constant scattering rate the random walk is a Poisson process. We verify our results with two iterative calculations, one using the properties of iterated convolution with a symmetric kernel and the other via direct calculation with an exponential step-length distribution. We present a novel demonstration, based on fluctuation theory of sums of random variables, that the first-passage probability as a function of the number

³Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

of peaks n in the alternating path is a step-length distribution-free combinatoric expression involving Catalan numbers. Counting paths with backscattering on the real half-line results in the same Catalan number coefficients as Dyck paths on the whole numbers. Including a separate forward-scattering Poisson process results in a combinatoric expression related to counting Motzkin paths. We therefore connect walks on the real line to discrete path combinatorics.

Keywords: diffusion in random media, fluctuation theorems, diffusion, stochastic processes

Contents

1. Introduction	2
2. Traditional solution of the Kubelka–Munk model	5
3. Distribution of path-lengths from the reflectance	6
4. Probability of first-passage by convolution	8
4.1. Iterations: building an alternating random walk	8
4.2. Dressing the skeleton: connection with combinatorics.....	10
5. Analytic calculation of the path distribution	11
6. First-passage events are distribution-free	13
References	17

1. Introduction

The perceived quality of a printed image is affected by three-dimensional light scattering in paper. The Kubelka–Munk [1] equations are most commonly used to estimate the effect of the medium on reflectivity. Kubelka and Munk provide a one-dimensional analytic solution for a model with two fluxes on the positive half-line, one moving upward into the medium in the positive direction and the other moving in the negative direction, downward, back toward the interface at the origin. While their approach is well-suited to analysis of reflectivity of layered media, it lacks the lateral scattering important to print quality. A light ray entering a three-dimensional scattering medium executes a random walk through a series of scattering events and exits at a different point. On a printed image this results in ‘optical gain’ as light that would enter at a printed point and exit nearby is absorbed by the print. The Yule–Nielsen [2] model extends Kubelka–Munk to describe optical gain. While its physical basis is not completely established, the Yule–Nielsen is an improved print-quality model. The impetus of this work is a more detailed understanding of scattering in one dimension as a step toward a true physically motivated three-dimensional scattering model with similar complexity to Kubelka–Munk.

J. Stat. Mech. (2020) 063404

The ubiquitous applications of random walks have inspired a large literature. The monograph by Rudnick [3] derives statistical properties of random walks and applies the theory to polymers and statistical mechanical systems. Redner's book [4] is devoted to first-passage processes. Depending on the problem, these authors switch between continuous and lattice models. The book by Spitzer [5], one of the pioneers in fluctuation theory, primarily addresses walks on a lattice. Philips–Invernizzi and Caze [6] provide a structured bibliographic review of the results obtained during the previous century.

Schwarzschild [7] investigated the temperature distribution within a stellar atmosphere using radiative equilibrium. Chandrasekhar's definitive theory for radiative transfer [8, 9] was developed in the context of plane-parallel layers of stellar and planetary atmospheres. The solutions of radiation transport equations have a variety of applications from neutron diffusion, optical tomography [10, 11], spreading of infra-red and visible light in the atmosphere to the quality of prints on paper.

The simplest solution of the radiative transfer equation is for a one-dimensional flux traveling perpendicular to a plane-parallel layer of an absorbing and scattering medium with isotropic radiation intensity over the forward and backward hemispheres. One-dimensional radiative transfer can be solved by the two-flux approximation proposed independently by Schuster [12] and Schwarzschild [7] in astronomy. Gate [13] examined the relation between Kubelka–Munk and the exact solution of the radiative transport equation. More recently, Sandoval [14] generalized the Kubelka–Munk theory using radiative transfer theory.

Youngquist *et al* [15] applied the model in optical coherence tomography and Haney and van Wijk [16] in geology. Hébert and Becker [17] examine the relation between continuous and lattice versions of Kubelka–Munk in an effort to give a more physical interpretation of the two-flux model. Ballestra *et al* [18] compute the first-passage probability density function using an integral representation obtained solving a system of convolution equations in the context of financial mathematics.

Simon and Trachsler [19] give an explicit expression for the reflectance. They suggest that the scattering problem can be treated as a Markov chain involving Narayana polynomials. But this Markov chain does not provide a solution of a first-passage problem that fits the reflectance calculated with the Kubelka–Munk equations. Then using the compositional optical reflectance and transmittance properties for multilayer specimens, they determine a generating function and find a solution as variants of Chebyshev polynomials. This is an alternative interpretation of the hyperbolic functions of the classical solution of Kubelka–Munk equations.

Monte Carlo simulations of random walks and first-passage process are commonly used to solve radiation transfer problems. Even in one dimension, Monte Carlo simulations run into the problem that the mean number of scattering events before first-passage is infinite. Jacques [20] developed code for scattering in multiple tissue layers. Doering *et al* [21] showed that the transmission time probability density has a long tail and anomalous moments due to multiple scattering events. Antal and Redner [22] study the first-passage of a random walk in an interval with a bounded uniformly distributed step length. They observe non-diffusive effects that persist when the maximum step-length is small, especially if the starting point is close to one edge.

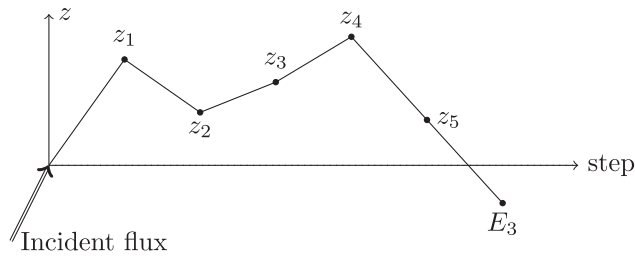


Figure 1. The flux enters at step 0 traveling in the positive, or upward, direction. After several scattering events, the path leaves the medium.

Wuttke [23] recursively expands the equations of Darwin [24] and Hamilton [25], or equivalently, the Kubelka–Munk equations. He identifies Catalan numbers in the recurrence probability as a function of the number of backscattering events. Wuttke defines a zigzag walk as an alternating walk where the random character comes from exponentially distributed step lengths between scattering events. As we demonstrate, it is not necessary to specify the distribution since his result is independent of the distribution of step-length distribution.

Scattering with a constant rate is a Poisson process. It can be analyzed by solving differential equations, as illustrated in section 2. Alternatively, the process can be treated as a random walk due to a sequence of scattering events with an exponentially distributed spacing. The first-passage process is a probabilistic process by which a fluctuating quantity reaches a threshold for the first time.

One of our goals is to describe the Kubelka–Munk solution directly in terms of the statistics of the number of peaks n and the path-length λ of rays. In one dimension the number of reflections before first-passage is $2n - 1$, so the scattering order, or number of reflections, is odd.

We explore Kubelka–Munk using a mathematically equivalent one-dimensional random walk model. Rays entering the medium from above is a more natural perspective, but to make contact with the random walk language, we use rays entering upwards from the negative z -axis. The scattering medium is on the positive z -axis. The incident flux enters the medium moving upwards in the positive direction from negative z . An example path is illustrated in figure 1. Backscattering or reflection occurs at z_1 , z_2 and z_4 . The step-lengths between scattering events are drawn from a random distribution. We include an additional independent Poisson process for forward scattering with a rate S_f . The example path undergoes forward scattering at z_3 and z_5 . Usually forward scattering is not included in one-dimensional scattering because it does not change the ray propagation. We include it here as a step towards future analysis of three-dimensional problems. In three dimensions a scattering event may change the lateral direction, but not the direction perpendicular to the surface. In this work, forward scattering and backscattering are assumed to be independent processes. We demonstrate that the recurrence probability given by Catalan numbers is independent of the step-length distribution. We provide another demonstration of the independence of step-length distribution using our formulation of the fluctuation theory introduced by Andersen [26, 27].

2. Traditional solution of the Kubelka–Munk model

This section illustrates the solution of the Kubelka–Munk [1] two-flux model as reviewed by Myrick *et al* [28]. Let us consider a homogenous layer with thickness d characterized by its absorption coefficient χ and its scattering coefficient S . In this layer, the incident irradiance I propagates in the positive direction and the reflected irradiance J propagates in the negative direction. Both I and J are functions of the depth z in the layer. Depth 0 corresponds to the layer’s boundary receiving the incident irradiance I_0 . Depth d indicates the other boundary. We consider, at an arbitrary depth z , a sub-layer with infinitesimal thickness dz . The effect of the material in a thin element dz on I and J is to:

- Decrease I by $I(S + \chi)dz$ (absorption and scattering).
- Decrease J by $J(S + \chi)dz$ (absorption and scattering).
- Increase I by $JSdz$ (scattered light from J reinforces I).
- Increase J by $ISdz$ (scattered light from I reinforces J).

On these assumptions we obtain the system of equations:

$$\begin{pmatrix} dI \\ dJ \end{pmatrix} = \begin{pmatrix} -(\chi + S) & S \\ -S & (\chi + S) \end{pmatrix} \begin{pmatrix} I \\ J \end{pmatrix} dz, \tag{1}$$

with solution

$$\begin{pmatrix} I(z) \\ J(z) \end{pmatrix} = \begin{pmatrix} 1 - \beta & 1 + \beta \\ 1 + \beta & 1 - \beta \end{pmatrix} \begin{pmatrix} Ae^{\kappa z} \\ Be^{-\kappa z} \end{pmatrix}, \tag{2}$$

where $\beta = \sqrt{\chi(\chi + 2S)^{-1}}$ and $\kappa = \sqrt{\chi(\chi + 2S)}$.

The coefficients A and B are determined by the boundary conditions at the two surfaces. After some elementary calculations reflectance R and transmittance T of a slab of thickness d are given by:

$$\begin{aligned} R &= \frac{J(0)}{I_0} = \frac{(1 - \beta^2)(e^{\kappa d} - e^{-\kappa d})}{(1 + \beta)^2 e^{\kappa d} - (1 - \beta)^2 e^{-\kappa d}} \\ R_0 &= \frac{Sd}{1 + Sd} \quad \text{for } \chi = 0 \\ T &= \frac{I(d)}{I_0} = \frac{4\beta}{(1 + \beta)^2 e^{\kappa d} - (1 - \beta)^2 e^{-\kappa d}} \\ T_0 &= \frac{1}{1 + Sd} \quad \text{for } \chi = 0. \end{aligned} \tag{3}$$

The reflectance of a very thick layer ($d \rightarrow \infty$) is:

$$R^\infty(S, \chi) = \frac{S + \chi}{S} - \sqrt{\left(\frac{S + \chi}{S}\right)^2 - 1}. \tag{4}$$

The Kubelka–Munk reflectance plays a major role in elucidating the connection to combinatorics and random walks.

3. Distribution of path-lengths from the reflectance

The fluxes I and J can be interpreted as ensembles of photons moving in the positive and negative directions. Each photon path alternates at scattering events between positive and negative steps with a sequence of step lengths $c_i \in \mathbf{c}$. Step i has length c_i chosen from a probability distribution $\rho(c_i)$ according to a Poisson process with rate S where

$$\rho(c) = \Phi(c)S e^{-Sc}. \tag{5}$$

Φ is the Heaviside step function. First, in the absence of absorption, the photon exits the medium after travelling a distance λ , the path-length. The probability distribution of λ for the random walk is $P(\lambda; S)$.

Each photon is absorbed at a rate χ per unit length. Therefore, its path is weighted according to the Beer–Lambert law [29, 30] by the absorption factor $e^{-\chi\lambda}$. The reflectance is therefore the Laplace transform of the path-length distribution and the inverse Laplace transform of R^∞ leads to the path distribution:

$$R^\infty(S, \chi) = \mathcal{L}_\lambda \{P(\lambda; S)\}(\chi) = \int_0^\infty P(\lambda; S) e^{-\chi\lambda} d\lambda \tag{6}$$

$$P(\lambda; S) = \mathcal{L}_\chi^{-1} \{R^\infty(S, \chi)\}(\lambda).$$

We calculate the inverse Laplace transform of the reflectance by first expanding equation (4) in the scattering order

$$R^\infty(S, \chi) = \sum_{n=1}^\infty \frac{C_{n-1}}{2^{2n-1}} \left(\frac{S}{S+\chi}\right)^{2n-1} \tag{7}$$

$$= \frac{1}{2} \frac{S}{S+\chi} C \left(\left(\frac{1}{2} \frac{S}{S+\chi}\right)^2 \right),$$

where $C(x)$ is the generating function

$$C(x) = \sum_{n=0}^\infty C_n x^n = \frac{1 - \sqrt{1 - 4x}}{2x} \tag{8}$$

of the Catalan numbers $C_n = (2n)!(n!(n+1)!)^{-1}$. We identify, term-by-term, the path-length distribution of a random walk model with the inverse Laplace transform of $R^\infty(S, \chi)$ using

$$\mathcal{L}_\chi^{-1} \left(\frac{S}{S+\chi}\right)^{2n-1} = \frac{S^{2n-1} \lambda^{2n-2} e^{-S\lambda}}{(2n-2)!}. \tag{9}$$

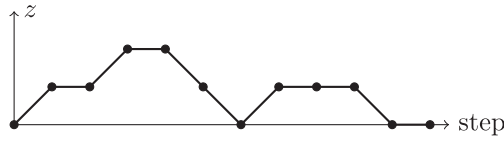


Figure 2. A typical Motzkin path with eleven steps.

The distribution of λ and n derived from $R^\infty(S, \chi)$ is

$$P(\lambda, n) = \frac{1}{\lambda} C_{n-1} \left(\frac{S\lambda}{2}\right)^{2n-1} \frac{e^{-S\lambda}}{(2n-2)!} \tag{10}$$

$$= \frac{1}{\lambda} \left(\frac{S\lambda}{2}\right)^{2n-1} \frac{e^{-S\lambda}}{n!(2n-1)!}. \tag{11}$$

The random character of an alternating walk comes from the step-length distribution between scattering events, which is not necessarily an exponential distribution. We will demonstrate in the paper that certain statistics of alternating walks is independent of the step-length distribution, i.e., is distribution-free.

With no loss of generality we can include n_f forward scattering events at a rate S_f thus providing a persistent random walk model compatible with Kubelka–Munk. Therefore the joint probability of the three random variables λ , n and n_f is given by

$$P(\lambda, n, n_f) = \frac{C_{n-1} S^{2n-1} S_f^{n_f} \lambda^{2n-2+n_f}}{2^{2n-1} (2n-2)! n_f!} e^{-(S+S_f)\lambda}, \tag{12}$$

and the joint probability for n and n_f is then given by

$$P(n, n_f) = \int_0^\infty P(\lambda, n, n_f) d\lambda$$

$$= \frac{1}{2^{2n-1}} \frac{[n_f + 2(n-1)]!}{n_f! n! (n-1)!} \frac{S_f^{n_f} S^{2n-1}}{(S_f + S)^{n_f+2n-1}}. \tag{13}$$

The combinatorial factor in equation (13) is $T(n_f + 2n - 2, n - 1)$ where

$$T(n, k) = \frac{n!}{k! (k+1)! (n-2k)!} \tag{14}$$

is the ‘Triangular array of Motzkin polynomial coefficients’. A Motzkin path is a path on an integer plane where each step is one of an up step (1, 1), a level step (1, 0) or a down step (1, -1). According to the OEIS A055151 [31] $T(n, k)$ is the number of Motzkin paths of length n with k up steps. The path shown in figure 2 has length $n = 11$ with $k = 3$ up steps.

4. Probability of first-passage by convolution

4.1. Iterations: building an alternating random walk

The previous results are derived from the inverse Laplace transform of the reflectance. We now want to study the trajectory distribution for a general path including both backscattering and forward scattering. We assume that forward scattering and backscattering are independent processes. We start with an alternating walk with only backscattering. The randomness comes only from the step length distribution. The scattering event locations form an alternating sequence of maxima and minima (peaks and valleys). First-passage necessarily occurs before a valley where the flux changes from the negative direction to the positive direction. Therefore, we study first-passage in the alternating walk by marginalizing over peaks. Finally, we introduce the forward scattering process. The alternating walk serves as a skeleton of the general walk and we then ‘dress the skeleton’ by adding forward scattering.

An alternating random walk, starts at 0 and moves upwards in the positive direction until it eventually backscatters downwards towards the origin. The walk continues through a sequence of upwards motion, reflecting at a peak to the negative direction, and subsequent reflection at a valley back to the positive direction. If a downwards step does not reverse direction before it reaches the negative half-space, we say it ‘left the medium’ and the trajectory ends. A trajectory that leaves the medium after n peaks is subject to $(2n - 1)$ backscattering events and is labeled by the index n .

Consider now the flux of all possible trajectories starting at $z = 0$ and beginning in the positive direction with step length c distributed according to equation (5).

Marginalizing over the first peak height gives the probability distribution $P_1(z_1)$ of the height z_1 of the first valley. If z_1 is negative, the trajectory ends at that point. $P_1(z_1)$ serves as a convolution kernel for moving from one valley to the next with the following properties:

$$\int_{-\infty}^{\infty} P_1(z_1) dz_1 = 1 \quad (15)$$

$$P_1(z_1) = P_1(-z_1).$$

P_1 is symmetric in its argument because the upward and downward steps are drawn from the same distribution.

The trajectory reflects back towards the origin after the n th peak at height z_n^P as shown in figure 3. It reflects next time in the positive direction at the n th valley at height z_n . The difference, $z_n - z_{n-1}$, in height of the valleys is the difference of the upward and downward step lengths. If $z_n < 0$, then we say the trajectory left the medium. The probability distribution for the height z_2 of the valley at the second reflection is

$$P_2(z_2) = \int_0^{\infty} P_1(z_1) P_1(z_2 - z_1) dz_1 \quad (16)$$

$$\int_{-\infty}^{\infty} P_2(z_2) dz_2 = \int_0^{\infty} P_1(z_1) dz_1 = \frac{1}{2}.$$

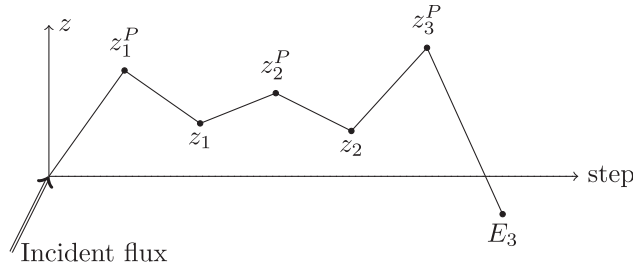


Figure 3. The flux enters at step 0 traveling in the positive, or upwards, direction. After several scattering events, the path can leave the medium. The figure shows an alternating path with first-passage at step six after the third peak.

This is a convolution of a symmetric function with itself. The symmetry of P_1 implies that, independent of the step-length distribution, half the trajectories leave the medium and half remain in the medium at the first valley. This simple case suggests that first passage could be independent of the step-length distribution.

The probability distribution for the location of the following valleys $n = 3, 4, \dots$ are given by iterative convolutions:

$$P_{n+1}(z_{n+1}) = \int_0^\infty P_n(z_n)P_1(z_{n+1} - z_n)dz_n. \tag{17}$$

Only the trajectories that stay in the medium at step n , are transferred to step $n + 1$, giving:

$$\int_{-\infty}^\infty P_{n+1}(z_{n+1}) dz_{n+1} = \int_0^\infty P_n(z_n) dz_n. \tag{18}$$

The probability of not escaping the medium after n peaks is

$$\int_{-\infty}^\infty \int_0^\infty P_n(z_n) P_1(z_{n+1} - z_n) dz_n dz_{n+1} = P^+(n) \tag{19}$$

and the probability of first-passage after n peaks is

$$\int_{-\infty}^\infty \int_{-\infty}^0 P_n(z_n) P_1(z_{n+1} - z_n) dz_n dz_{n+1} = P^f(n). \tag{20}$$

with the property

$$\sum_{m=1}^n \int_{-\infty}^\infty \int_{-\infty}^0 P_m(z_m) P_1(z_{m+1} - z_m) dz_m dz_{m+1} + \dots \tag{21}$$

$$\int_{-\infty}^\infty \int_0^\infty P_n(z_n) P_1(z_{n+1} - z_n) dz_n dz_{n+1} = 1,$$

since the sum of the fractions that leave the medium after each peak plus the fraction that remain in the medium must add to 1. Solving the iteration defined by equations (15) to (20) requires an additional property. Extensive Monte Carlo results with a variety

of step-length distributions led us to believe that $P^+(n) = (2n - 1)P^f(n)$, independent of the choice of step-length distribution. We solve the iteration inductively using this property giving results consistent with the Laplace transform of the reflectance in the previous section:

$$P^+(n) = \prod_{m=0}^{n-1} \frac{1 + 2m}{2 + 2m} = \frac{2n - 1}{2^{2n-1}} C(n - 1) \tag{22}$$

$$P^f(n) = \frac{1}{2^{2n-1}} C(n - 1).$$

These results will be confirmed by an analytic calculation of the path distribution in the case of an exponentially distributed path-length, consistent with Kubelka–Munk, in section 5. The case for a general path-length distribution is shown by a combinatorial method in section 6 equations (40) and (41).

4.2. Dressing the skeleton: connection with combinatorics

The enumeration of lattice paths is a topic in combinatorics, closely related to the study of random walks in probability theory. The ubiquitous presence of Catalan numbers in the joint distribution function $P(\lambda, n, n_f; S, S_f)$ suggests a connection with combinatorics. One approach to explain this connection is discretization. Since the statistics of first-passage is independent of the distribution of the path-length, to discretize the path we just have to integrate over λ . We expect this transition to a lattice model to reproduce the analytical result obtained in section 6.

The joint probability function is the product of the marginal probability times the conditional probability:

$$P(n, n_f; S, S_f) = P(n_f | n; S, S_f) P(n; S, S_f). \tag{23}$$

The goal is to create new steps by randomly filling the $2n$ segments of the skeleton with n_f forward random scattering events and to calculate the resulting distribution. Ultimately we want to find the related conditional probability $P(n | n_f)$. There are $2n - 1$ reflections and $m_s = 2n + n_f$ steps.

With our notation we have paths from $(2, 0)$ to $2(n, n_f)$ with the constraint that $n \geq 1$. Therefore the number of paths N_C is given by:

$$N_C = \binom{2n + n_f - 2}{2n - 2}. \tag{24}$$

In terms of the number of scattering event m_s the number of paths is

$$N_C(m_s, n) = \frac{(m_s - 2)!}{(2n - 2)!(m_s - 2n)!}. \tag{25}$$

We can now write the conditional probability of m_s at constant n with $r = \frac{S}{S+S_f}$

$$P(m_s | n; r) = \frac{(m_s - 2)!}{(2n - 2)!(m_s - 2n)!} r^{2n-1} (1 - r)^{m_s-2n}. \tag{26}$$

This result has been confirmed by extensive Monte Carlo calculations and is also obtained by recursion in section 5. Then the joint probability of n and n_f is given by

$$\begin{aligned}
 P(n, n_f; r) &= \frac{[n_f + 2(n - 1)]!}{n_f! n! (n - 1)!} \left(\frac{r}{2}\right)^{2n-1} (1 - r)^{n_f} \\
 &= \frac{(n_f + 2(n - 1))!}{2^{2n-1} n_f! (n - 1)! (n)!} \frac{S^{2n-1} (S_f)^{n_f}}{(S + S_f + \chi)^{n_f + 2n - 1}}.
 \end{aligned}
 \tag{27}$$

This is the same as equation (13). This confirms the combinatorial nature (lattice paths enumeration) of the discrete form of the Kubelka–Munk equation.

5. Analytic calculation of the path distribution

In this section, explicit integration over the path gives the distribution of path-length and number of scattering events as a function of depth into the medium. We include forward scattering, reflection and absorption as independent processes acting on an ensemble of light rays. We show directly the connection to equation (4), as opposed to the two-step calculation in section 4. We derive information about the distribution of the number of scattering events and the path-length when the ray ultimately escapes. The current state of a ray at a scattering event includes the direction, the total prior path-length λ , the prior number of peaks n and the position z . This will serve as a foundation for future exploration of higher dimensions and for the effect of inhomogeneities in the medium that are important for print quality and medical imaging.

Consider a statistical ensemble of rays of light moving in a one-dimensional diffusive medium on the positive axis. Again, the ray starts at the origin moving upward in the positive direction and after multiple reflections, escapes the media to the negative axis. A ray traveling through the medium is reflected at a rate of S reflections per unit length. The Poisson probability density for reflection of a ray after traveling distance c is given by equation (5). Later we will add forward scattering at a rate S_f and absorption at a rate χ per unit length as independent Poisson processes based on the path-length.

The probability density for an upward reflection at a valley at height z with current path-length λ at the valley after n peaks is $P_n(\lambda, z)$. The probability density that the ray escapes with total path-length λ in the scattering medium, after the n th peak is $E_n(\lambda)$. The probability densities for the peaks and valleys are not normalized because the ray may escape on a previous step.

The initial path-length is 0 and the initial position is the origin. The first step is into the medium, so the initial escape probability $E_0 = 0$.

$$\begin{aligned}
 P_0(\lambda, z) &= \delta(\lambda) \delta(z) \\
 E_0(\lambda) &= 0.
 \end{aligned}
 \tag{28}$$

The first peak integral simply involves satisfying the delta functions. Similarly, the first valley integral uses the fact that λ is the total path-length. The joint probability

distribution of the height and path-length for the first peak is

$$P_1^P(\lambda, z^P) = \int_0^\lambda \int_0^\infty P_0(\lambda', z') \delta(\lambda - (\lambda' + z^P - z')) \rho(z^P - z') d\lambda' dz' \quad (29)$$

$$= \delta(\lambda - z^P) \rho(z^P).$$

Integrating over the first peak gives the joint distribution of the path-length λ and height z_1 of the first valley

$$P_1(\lambda, z_1) = \int_0^\infty dz_1^P \delta(\lambda - 2z_1^P + z_1) S e^{-S z_1^P} \rho(z_1^P - z_1) \quad (30)$$

$$= \frac{1}{2} \Phi(z_1) \Phi(\lambda) \Phi(\lambda - z_1) S^2 e^{-S\lambda}.$$

The probability that the ray escapes is given by dropping the constraint that z is positive and integrating over the negative half space:

$$E_1(\lambda) = \frac{1}{2} \Phi(\lambda) \int_{-\infty}^0 S^2 e^{-S(\lambda-z)} dz = \frac{1}{2} \Phi(\lambda) S e^{-S\lambda}. \quad (31)$$

Integrating over λ gives the expected escape rate of $\frac{1}{2}$ after the first peak.

The probability distribution of the n th peak and valley are found by iteratively calculating the convolution using an upward step with length c' and a downward step with length c giving

$$P_n^P(\lambda, z_n^P) = \int_0^{z_n^P} P_{n-1}(\lambda - c', z_n^P - c') S e^{-S c'} dc' \quad (32)$$

$$P_n(\lambda, z_n) = \int_z^\infty P_n^P(\lambda - c, z_n + c) S e^{-S c} dc$$

$$P_n(\lambda, z_n) = \int_0^\infty \int_0^{c+z_n} P_{n-1}(\lambda - c' - c, z_n + c - c') S^2 e^{-S(c+c')} dc' dc.$$

There are constraints on the integrals in iterating from one valley to the next. The peaks are higher than the preceding and following valleys, i.e., $z_n^P \geq z_n$ and $z_n^P \geq z_{n-1}$, as shown in figure 3.

Iterating from the probability distribution at one valley to that at the next valley, we obtain:

$$P_n(\lambda, z_n) = S^{2n} e^{-S(\lambda)} \Phi(\lambda - z_n) F_n(\lambda, z_n). \quad (33)$$

Here F_n is the total volume of the space of allowed configurations of the $2n - 2$ step path from the origin to a valley at the point z with path-length λ . The same form, an exponential times the volume of path configurations, will apply in higher dimensions and more complicated geometries. Monte Carlo methods can be applied to measure the path configuration volume in these situations.

$$F_n(\lambda, z) = \frac{(\lambda - z)^{n-1} (\lambda + z)^{n-2} (\lambda - z + 2nz)^{n-1}}{2^{2n-2} (n-1)! (n)!}. \quad (34)$$

Integrating over the negative half-space gives the escape probability density as a function of upward path-length and the total escape probability after the n th peak:

$$E_n(\lambda) = \int_{-\infty}^0 P_n\left(\frac{\lambda}{2}, z'\right) dz' = S\left(S\frac{\lambda}{2}\right)^{2(n-1)} \frac{e^{-S\lambda}}{(n-1)!(n)!}. \tag{35}$$

Integrating over λ gives the escape probability after n peaks in terms of Catalan numbers. The result is independent of the scattering rate and the formula is identical to equation (3) with $\chi = 0$:

$$E_n = \int_0^\infty E_n(\lambda) d\lambda = \frac{1}{2^{2n-1}} C_{n-1}. \tag{36}$$

We can add in absorption because we have the distribution of the path-length. The attenuation simply adds χ to S in the exponent:

$$E_n^\chi(\lambda) = S(S\lambda)^{2n-1} \frac{e^{-(S+\chi)\lambda}}{(n-1)!(n)!}. \tag{37}$$

Integrating over the path-length again yields equation (3)

$$\begin{aligned} E_n^\chi &= \int_0^\infty E_n^\chi(\lambda) d\lambda = \left(\frac{S}{S+\chi}\right)^{2n-1} \frac{(2n-1)!}{2^{2n-1}(n-1)!(n)!} \\ &= \left(\frac{S}{S+\chi}\right)^{2n-1} \frac{C_{n-1}}{2^{2n-1}}. \end{aligned} \tag{38}$$

The joint probability for n and n_f is obtained by adding an independent Poisson process for forward scattering

$$\begin{aligned} P(n_f, n) &= \int_0^\infty P_{\text{Pois}}(n_f | S_f \lambda) E_n^\chi(\lambda) d\lambda \\ &= \frac{(n_f + 2(n-1))!}{2^{2n-1} n_f! (n-1)!(n)!} \frac{S^{2n-1} (S_f)^{n_f}}{(S + S_f + \chi)^{n_f + 2n - 1}}. \end{aligned} \tag{39}$$

This result is identical to equations (13) and (27).

6. First-passage events are distribution-free

This section extends the fluctuation theory of Andersen [27] to alternating walks with independent and identically distributed step-lengths. We prove first-passage is distribution-free by examining first-passage properties of the finite set of alternating walks generated from permutations of a set of $2n$ step-lengths. Surprisingly, the probability of first-passage at each valley for this set of walks is independent of the set of step-lengths.

Consider the sample space \mathcal{A}_n of alternating walks with $2n$ steps. In this section we use the capitalized term ‘Event’ in the probabilistic sense. An event is a subset of

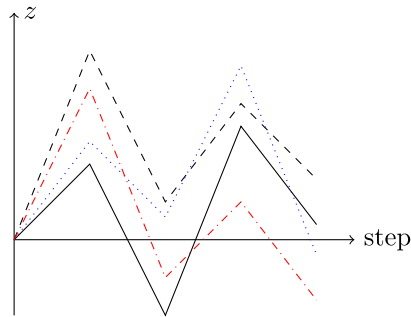


Figure 4. Four examples of the 24 alternating walks generated from permutations of an example set of four lengths $\mathbf{c}_2 = \{1.0, 1.3, 2.0, 2.5\}$.

the sample space of \mathcal{A}_n . The indicator function \mathbb{I}_E of an event is 1 for a walk that is an element of the event and 0 otherwise. The probability of an event is given by the average over \mathcal{A}_n of the indicator function \mathbb{I}_E of the event. First-passage events are the subsets \mathcal{F}_m of \mathcal{A}_n that are positive before peak m and first become negative at step $2m$ after peak m . It is possible that a walk never becomes negative, so we add the event \mathcal{F}_0 of walks that are positive for the first $2n$ steps. The set of $n + 1$ first-passage events, one at each valley plus one if the walk stays positive, is a partition of the sample space.

The location of an alternating walk alternates between peaks and valleys as shown in figure 3. First-passage for an alternating walk with n peaks occurs at a valley where the step number $2m$ is even. The positions of the peaks and valleys are given by repeated convolution with the step-length distribution as illustrated in sections 4 and 5. The set of $2n$ step-lengths \mathbf{c}_n inside the integrand of the iterated convolution is an element of the set \mathcal{C}_n of all sets of $2n$ lengths. Because the step-lengths are independent and identically distributed, the joint probability density in the integrand is symmetric under permutations of the $2n$ step-lengths \mathbf{c}_n in the integrand. The average of the indicator function \mathbb{I}_E over \mathcal{A}_n can therefore be evaluated by first averaging over the finite set of walks generated by permutations of \mathbf{c}_n . Figure 4 illustrates, for example, four of the 24 walks generated from permutations of a particular \mathbf{c}_2 . The surprising result, shown below, is that this average over a finite set of walks is independent of the set \mathbf{c}_n .

Consider the finite set of walks generated by permutations of \mathbf{c}_n . Each of the alternating walks in $\mathcal{A}_n(\mathbf{c}_n)$ generated by permutations of \mathbf{c}_n has the same weight in the integral. We will show that the cardinality of $\mathcal{F}_m(\mathbf{c}_n)$ is independent of the set of $2n$ step lengths \mathbf{c}_n almost everywhere, as long as no sub-walk returns exactly to the origin.

We analyze the changes in event membership as we modify step-lengths in \mathbf{c}_n . The boundary subset \mathcal{B}_n of \mathcal{C}_n of measure zero, consists of those sets of step-lengths \mathbf{c}_n where some sub-walk constructed from \mathbf{c}_n returns to the origin. For a walk constructed from a set of lengths not in the boundary set \mathcal{B}_n , let δ be the closest approach to the origin of any non-empty sub-walk. At step $2m > 0$ the absolute value of the position must be at least δ . If $\delta > 0$ then any step-length change with magnitude smaller than δ will not

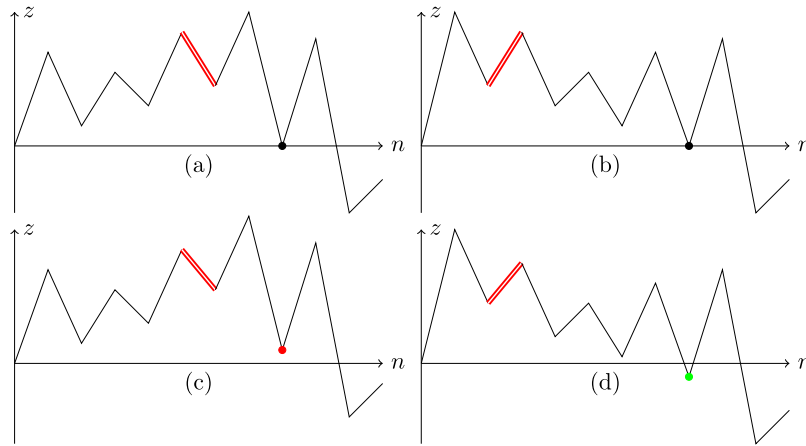


Figure 5. Example of pairing critical walks for the event ‘First-passage on step 8’. Critical walk (a) is paired with critical walk (b) with the sub-walk for the first eight steps reversed. When the sixth step length is shortened, (c) shows walk (a) leaving the event and (d) shows walk (b) joining the event.

cause any change in first-passage event membership. First-passage event cardinality is therefore locally constant, and so is constant within connected subsets of $\mathcal{C}_n - \mathcal{B}_n$. The set $\mathcal{C}_n - \mathcal{B}_n$ is not connected, so we must examine changes in event cardinality when crossing a boundary.

Any set of lengths can be reached by changing one length at a time, and so if cardinality of an event never changes when crossing the boundary \mathcal{B}_n , then the event cardinality is constant in $\mathcal{C}_n - \mathcal{B}_n$. The sequences of lengths in \mathcal{B}_n are the only elements of \mathcal{C}_n where walks can leave or enter first-passage events with small changes to a length. Consider changing a length in a walk w_m in \mathcal{F}_m with first-passage following the m th peak. It can move across the boundary set \mathcal{B}_n . As it crosses the boundary point the membership of w_m in \mathcal{F}_m can change. The set of walks generated from permutations or an overall sign change of the steps in the sub-walk w_m^* also return to the origin. The problem is to show cancellation of the changes in event membership among the walks in this set as lengths are changed.

The key step in our approach is to consider uniquely defined pairs of walks w^* and w'^* from the boundary set. Walk w^* begins with a positive critical sub-walk w_m^* as in figure 5(a), but returns precisely to zero at step $2m$ as illustrated by the dot at step 8. The paired walk w'^* in figure 5(b) is identical after step $2m$ but begins with the time-reversed sub-walk $w_m'^*$. The sets of positions of the paired walks are the same, but they are in the reverse order for the first $2m$ steps. The step lengths in $w_m'^*$ are the same as the step lengths in w_m^* , but the first $2m$ steps occur in the reverse order and with the opposite sign. All the critical walks generated from $\mathbf{c}_n^* \in \mathcal{B}_n$ that could change the cardinality of \mathcal{F}_m can be uniquely paired in this way.

For the first-passage problem, \mathcal{F}_m consists of walks that are positive for the first $2m - 1$ steps and negative on step $2m$. The boundary of \mathcal{F}_m consists of walks that are non-negative for the first $2m - 1$ steps but return exactly to zero at one of the first $2m$ steps. Walks that return precisely to the origin on step $2m$ and are positive before

that step for steps $1, \dots, (2m - 1)$ will enter or leave the event with a small change in a length. Similarly, walks that pass zero on step $2m$ but return precisely to the origin on step $2j < 2m$ change membership in \mathcal{F}_m with small length changes.

Suppose a walk from the boundary of \mathcal{F}_m begins with such a $2j$ -step sub-walk w_j^* which is positive for $k < 2j$ and returns to 0 at step $2j$. Consider changing a length $c_k \in w_j^*$ through the critical value c_k^* . A small decrease in a length that is a downward step in w_j^* will cause the walk to leave \mathcal{F}_j and join \mathcal{F}_m as shown in figure 5(c) while increasing the same step-length will cause the walk to be in \mathcal{F}_j . Similarly, a small decrease in a length that is an upward step in w_j^* will cause the walk to be in \mathcal{F}_j as shown in figure 5(d) while increasing the same step will cause the walk leave \mathcal{F}_j and join \mathcal{F}_m . Alternatively, if the walk touches zero at step m then increasing a downward length will cause it to be in \mathcal{F}_m while the paired walk will leave \mathcal{F}_m .

Figures 5(a) and (b) illustrates paired critical walks that touch the origin at step eight. When the step illustrated by the double red line is shortened, the trajectory in figure 5(c) leaves the event of passage at step eight and $z_8 > 0$ as indicated by red dot above the axis. The paired walk in figure 5(d) enters the event as the dot moves below the line indicating $z_8 < 0$.

If a step of length c_k^* in sub-walk w_j^* has, say, a negative sign, then it is in $w_j'^*$ with a positive sign. When c_k^* is changed to a lower value then walk w' is in \mathcal{F}_m and w is not. Similarly, when c_k is changed to a higher value, then walk w is in \mathcal{F}_m and w' is not. Thus as c_k is changed and the set of lengths crosses \mathcal{B}_n one of each pair of walks leaves the event \mathcal{F}_m and the other joins the event. Although the pair of walks switch which one is in event \mathcal{F}_m , the total contribution of the two walks to the cardinality of the event is the same. The cardinality of \mathcal{F}_m is therefore unchanged by crossing \mathcal{B}_n . Examining the complete set of pairs of walks with time-reversed sub-walks thus shows that the cardinality of first-passage events, and thus the probability of first-passage at step $2m$ is independent of the set of n real lengths, except for the boundary set \mathcal{B}_n of measure 0 in \mathcal{C}_n . Now averaging this invariant probability over all sets of lengths in \mathcal{C}_n gives the result that the distribution of first-passage step for alternating walks is step-length distribution-free.

The sub-walk must have even length for alternating walks so that walks beginning with w and w' are both alternating walks in our sample space. First-passage always occurs on an even number step for alternating walks, so that is not a problem here. In other types of events on \mathcal{A}_n the requirement that the sub-walks must have even length is important. The argument for first-passage statistics for symmetric walks carries through the same as for alternating walks with the exception that the location of the first-passage, and the length of the relevant sub-walks, need not be even.

To calculate first-passage probabilities for an alternating walk, any set of lengths suffices. Make a convenient choice like any subset of integer powers of 2 where each length is larger than the sum of all smaller lengths. First-passage of an alternating walk occurs only in a valley, i.e., on even numbered steps $m = 2m_p$. Given a set of $2n$ lengths selected from integer powers of 2, we show below that fraction of walks generated from these lengths with first-passage at step n is

$$P_n^f = \frac{C_n}{2^{2n-1}} \quad (40)$$

and the fraction that remains nonnegative is

$$P_n^+ = (2n - 1) \frac{C_{n-1}}{2^{2n-1}}. \quad (41)$$

The calculation will proceed by induction on n . The theorem is true for $n = 1$ because the probability of first passage in the first valley is $\frac{1}{2}$, as is the probability that the walk stays positive. Either the first step upward is larger or smaller than the second step.

Suppose the theorem is true for all $m_p \leq n$. Consider alternating walks generated from a set $\mathbf{c}_{2(n+1)}$ of $2(n+1)$ lengths selected from integer powers of 2. Divide the alternating walks generated into a complete set of disjoint subsets where all walks in a subset have the same last two steps. For each of these subsets, the first $2n$ steps are all the permutation of the same set of lengths. By induction, the fraction of these in \mathcal{F}_{2m_p} for $m_p \leq n$ is given by the theorem. Similarly, the fraction that stay positive until the last step is given by the theorem.

First passage can occur after the peak $n+1$ only if the walk stayed positive for the first n valleys and step $2n$ is the largest element of the set. The probability of this is

$$P_{n+1}^f = \frac{P_n^+}{2n+2}. \quad (42)$$

The walk will stay positive only if it is positive for the first n valleys and it does not have a first passage at valley $n+1$, so

$$P_{n+1}^+ = p_n^+ - P_{n+1}^f$$

in agreement with equation (41). The theorem is true for $n+1$, and so is proved by induction. This combinatorial approach finally proves the assertion we postulated based on the results of Monte Carlo simulations.

References

- [1] Kubelka P and Munk F 1931 An article on optics of paint layers *Z. Tech. Phys.* **12** 593–601
- [2] Yule J A C and Nielsen W J 1951 The penetration of light into paper and its effect on halftone reproduction *Proc. TAGA* vol 3 65–76
- [3] Rudnick J and George G 2004 *Elements of the Random Walk: An Introduction for Advanced Students and Researchers* (Cambridge: Cambridge University Press)
- [4] Redner S 2001 *A Guide to First-Passage Processes* (Cambridge: Cambridge University Press)
- [5] Spitzer F 1964 *Principles of Random Walk* Graduate Texts in Mathematics vol 34 (New York: Springer)
- [6] Philips-Invernizzi B, Dupont D and Caze C 2001 Bibliographical review for reflectance of diffusing media *Opt. Eng.* **40** 1082–92
- [7] Schwarzschild K 1906 On the equilibrium of the sun's atmosphere *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen* (Math.-Phys. Klasse vol 195) pp 41–53
- [8] Chandrasekhar S 1950 *Radiative Transfer* (Oxford: Clarendon)
- [9] Chandrasekhar S 1943 Stochastic problems in physics and astronomy *Rev. Mod. Phys.* **15** 1
- [10] Bonner R F, Nossal R, Havlin S and Weiss G H 1987 Model for photon migration in turbid biological media *J. Opt. Soc. Am. A* **4** 423–32

- [11] Nossal R, Kiefer J, Weiss G H, Bonner R, Taitelbaum H and Havlin S 1988 Photon migration in layered media *Appl. Opt.* **27** 3382–91
- [12] Schuster A 1905 Radiation through a foggy atmosphere *Astrophys. J.* **21** 1
- [13] Gate L F 1974 Comparison of the photon diffusion model and Kubelka–Munk equation with the exact solution of the radiative transport equation *Appl. Opt.* **13** 236–8
- [14] Sandoval C and Kim A D 2014 Deriving Kubelka–Munk theory from radiative transport *J. Opt. Soc. Am. A* **31** 628
- [15] Youngquist R C, Carr S and Davies D 1987 Optical coherence-domain reflectometry: a new optical evaluation technique *Opt. Lett.* **12** 158–60
- [16] Haney M M and van Wijk K 2007 Modified Kubelka–Munk equations for localized waves inside a layered medium *Phys. Rev. E* **75** 036601
- [17] Hébert M and Becker J-M 2008 Correspondence between continuous and discrete two-flux models for reflectance and transmittance of diffusing layers *J. Opt. A: Pure Appl. Opt.* **10** 035006
- [18] Ballestra L V, Pacelli G and Radi D 2016 A very efficient approach to compute the first-passage probability density function in a time-changed Brownian model: applications in finance *Physica A* **463** 330–44
- [19] Simon K and Trachsler B 2003 A random walk approach for light scattering in material *Discrete Random Walks, DRW'03* (Paris, France 2003) pp 289–300
- [20] Jacques S L 2010 Monte Carlo modeling of light transport in tissue (steady state and time of flight) *Optical-Thermal Response of Laser-Irradiated Tissue* (Berlin: Springer) pp 109–44
- [21] Doering C R, Ray T S and Lawrence Glasser M 1992 Long transmission times for transport through a weakly scattering slab *Phys. Rev. A* **45** 825–8
- [22] Antal T and Redner S 2006 Escape of a uniform random walk from an interval *J. Stat. Phys.* **123** 1129–44
- [23] Wuttke J 2014 The zig-zag walk with scattering and absorption on the real half line and in a lattice model *J. Phys. A: Math. Theor.* **47** 215203
- [24] Darwin C G 1922 XCII. The reflexion of x-rays from imperfect crystals *London, Edinburgh, and Dublin Phil. Mag. J. Sci.* **43** 800–29
- [25] Hamilton W C 1957 The effect of crystal shape and setting on secondary extinction *Acta Crystallogr.* **10** 629–34
- [26] Andersen E S 1962 The equivalence principle in the theory of fluctuations of sums of random variables *Colloquium on Combinatorial Methods in Probability Theory* (Aarhus, Denmark: Matematisk Institut Aarhus Universitet) pp 13–6
- [27] Sparre Andersen E 1953 On the fluctuations of sums of random variables *Math. Scand.* **1** 263–85
- [28] Myrick M L, Simcock M N, Baranowski M, Brooke H, Morgan S L and McCutcheon J N 2011 The Kubelka–Munk diffuse reflectance formula revisited *Appl. Spectrosc. Rev.* **46** 140–65
- [29] Beer A 1852 Bestimmung der absorption des rothen lichts in farbigen flussigkeiten *Ann. Phys.* **162** 78–88
- [30] Lambert J H 1760 *Photometria, sive de mensura et gradibus luminis. Colorum et Umbrae* (Augsberg: Eberhard Klett)
- [31] Sloane N J A *et al* 2019 *The On-Line Encyclopedia of Integer Sequences* (<https://oeis.org/A055151>)