A Stochastic Interpretation of Kubelka-Munk

Klaus Simon

Swiss Federal Laboratories for Materials Testing and Reserach, EMPA, St. Gall

Abstract

Understanding reflection is one of the key competences in graphic arts industry. A very popular approach was given by KUBELKA-MUNK [1] who derived a simple relationship between the scattering and absorption coefficients and the overall reflectance. In the course of time, the theory was extended by several authors, see [2], [3] for recent improvements. This paper presents an alternative approach which describes the behavior of light in matter as a random walk. In this respect KUBELKA-MUNK is closely correlated to both well-known stochastic theories and recent combinatorial research, in particular with catalan numbers.

1. Introduction

The original theory of KUBELKA-MUNK dates back to 1931, see[1]. A horizontal colorant layer with thickness D is considered. We suppose that the paint (material) is homogeneous. Then any light inside the layer travelling in any direction can be divided into its vertical and horizontal components. For simplification, the horizontal components are ignored. Therefore only two vertical fluxes of light have to be considered: a downward flux i and a upward flux j. Let i_x and j_x be the intensities of these two fluxes at the distance x, $x \leq D$, from the top surface. Then the KUBELKA-MUNK theory is based on the assumption that the fractional amount of light lost by absorption (scattering) between x and $x + \mathrm{d}\,x$ is given by $K\mathrm{d}\,x$ ($S\mathrm{d}\,x$), where K (S) is denoted by absorption (scattering) coefficient. Light absorbed between x and $x + \mathrm{d}\,x$ is lost, but light scattered from direction i is added to j and viceversa[1]. On these assumptions we obtain

$$\frac{\mathrm{d}\,i_x}{\mathrm{d}\,x} \;=\; (K+S)\cdot i_x - S\cdot j_x \tag{1}$$

and

$$\frac{\mathrm{d}\,j_x}{\mathrm{d}\,x} \;=\; S\cdot i_x - (K+S)\cdot j_x \tag{2}$$

which implies

$$\frac{\mathrm{d}\,r_x}{\mathrm{d}\,x} \;=\; S\cdot {r_x}^2 - 2\cdot(K+S)\cdot r_x + S, \tag{3}$$

where

$$r_x \;=\; \frac{j_x}{i_x}$$

is called the *reflectance ratio*. Now, the aim is the determination of the *reflectance*

$$R \;\stackrel{\text{def}}{=}\; r_0.$$

A first solution of equation (3) arises if the layer is so thick that further increases in thickness do not significantly change its reflectance, i.e.

$$\frac{\mathrm{d}\,r_x}{\mathrm{d}\,x} \;=\; 0.$$

In this situation we use R_∞ for the value of $(r_0)_{D\to\infty}$ and hence

$$R_\infty \;=\; 1 + \frac{K}{S} - \sqrt{\Big(1+\frac{K}{S}\Big)^2 - 1} \tag{4}$$

which is equivalent to

$$\frac{K}{S} \;=\; \frac{(1-R_\infty)^2}{2R_\infty}. \tag{5}$$

This was the original result of KUBELKA-MUNK [1]. Well-known variants are proposed by FOOTE [4], [5]

$$R_\infty \;=\; 1 + \frac{K}{S} - \sqrt{\frac{K^2}{S^2} + \frac{2K}{S}} \tag{6}$$

and by SAUNDERSON[6]

$$R_\infty \;=\; \frac{1}{1 + \frac{K}{S} + \sqrt{\frac{K^2}{S^2} + \frac{2K}{S}}}. \tag{7}$$

Later, KUBELKA solved the equation (3) on the additional assumption that the reflectance of background (bottom) is known. Let R_g be this fraction. Then the general solution is

$$R \;=\; \frac{1 - R_g\cdot(a - b\cdot\coth(bSD))}{a - R_g + b\cdot\coth(bSD)}, \tag{8}$$

where

$$a \;=\; 1 + K/S, \qquad b \;=\; \sqrt{a^2-1}$$

[1] note that K and S are functions of wavelength

and coth means the hyberbolic cotangent.

Up to now FRESNEL-effects of the interface between outside and inside of the colorant layer are ignored. For its top boundary SAUNDERSON [6, pp.728] found the following[2]: The relation between the reflection R' from a sample as measured by the spectrophotometer, and the "reflectance" R which the sample would have if measured in a transparent medium having the same index of refraction as the sample, can be expressed as

$$R' \;=\; \frac{k_1}{2} + (1-k_1)\cdot(1-k_2)\cdot\frac{R}{1-k_2R}. \qquad (9)$$

In this expression, k_1 is the fraction of the incident light which is reflected from the front surface of the sample, and can be found from FRESNEL's law for normal incidence for a change of index of refraction from 1 to n:

$$k_1 = \left[\frac{n-1}{n+1}\right]^2.$$

Of this fraction, only one-half is measured by the spectrophotometer, and only $(1-k_1)$ of the incident light enters the reflecting sample. The constant k_2 is the fraction of the light incident *diffusely* upon the surface of the sample from the inside which is reflected, so that the fraction $(1-k_2)$ emerges from the sample into the integrating sphere.

So far so good. Due to its simplicity, the theory of KUBELKA-MUNK is in common usage for industrial applications. However, this concept has also disadvantages, in particular, from a theoretical point of view. For instance, it seems to be a nontrivial task to integrate FRESNEL-effects as boundary condition into the differential equation (3) [3], [2]. Here we try a new approach for the foregoing phenomena: instead of treating fluxes of light, we model the travelling of a photon as a special random process similar to a *random walk*. Our intention is a clear separation of physical effects from the mathematical concept, a new understanding of the underlying mathematics, an easier interpretation of the involved parameter and a greater accuracy of the derived formulas. Furthermore, we hope that this model can be extended to higher dimensions.

2. Photon Motion and Random Walks

A discrete-time *birth-and-death process*[3] can be described as a sequence of random variables x_t, $t \in \mathbb{N}$, assuming the states $\ell = 0, 1, 2, \ldots$ with probability $P_{t,\ell}$. We suppose that the process starts at state 0 and epoch 0, hence $P_{0,0} = 1$. Direct transitions to state ℓ are only possible from state $\ell-1$ and $\ell+1$. The probability that such a transition takes place between epoch $t-1$ and t is noted by[4]

$$\lambda_{t-1,l-1} \;=\; \Pr(X_t = \ell \mid X_{t-1} = \ell - 1) \qquad (10)$$

and

$$\beta_{t-1,\ell+1} \;=\; \Pr(X_t = \ell \mid X_{t-1} = \ell + 1). \qquad (11)$$

Consequently, the process statisfies

$$\begin{aligned} P_{t,\ell} \;=\; & (1 - \lambda_{t-1,\ell} - \beta_{t-1,\ell}) \cdot P_{t-1,\ell} \\ & + \lambda_{t-1,\ell-1} \cdot P_{t-1,\ell-1} \\ & + \beta_{t-1,\ell+1} \cdot P_{t-1,\ell+1}. \end{aligned} \qquad (12)$$

The solution of recursion (12) clearly depends on $\lambda_{t,\ell}$ and $\beta_{t,\ell}$. A *random walk* is a discrete time birth-and-dead process with

$$\lambda_{t,\ell} \;=\; \lambda, \qquad \beta_{t,\ell} \;=\; \beta$$

and

$$\lambda + \beta \;=\; 1,$$

which simplifies (12) to

$$P_{t,\ell} \;=\; \lambda \cdot P_{t-1,\ell-1} + \beta \cdot P_{t-1,\ell+1}.$$

A *random walk with absorbing barriers at* $B_1 \le 0$ *and* $B_2 \ge 0$ is a random walk which stops at epoch t if

$$B_1 \le X_{t'} \le B_2 \quad \text{for } 0 \le t' < t$$

and

$$X_t \in \{\, B_1, B_2 \,\}.$$

Remark. A popular interpretation is given by FELLER[7, p. 342]: *Consider the familiar gambler who wins or loses a dollar with probabilities λ and β respectively. Let his initial capital be B_2 and let him play against an adversary with initial capital $|B_1|$. The game continues until the gambler's capital either is reduced to zero or has increased to $|B_1| + B_2$ that is, until one of the two player is ruined. We are interested in the probability of the gambler's ruin and the probability distribution of the duration of the game. This is the classical ruin problem.*

Next, we have to model the travelling of a photon as a random walk with absorbing barriers. First, we suppose that the photon moves in every discrete time step from the discrete level ℓ to $\ell+1$ or $\ell-1$. The entrance level is $\ell = 0$. The absorbing barrier

$$B_1 \;=\; -1$$

represents the outside of the colorant layer. The second barrier

$$B_2 \;=\; d + 1$$

[2] It is not known to the authors that a SAUNDERSON correction for both boundaries or a corresponding improvement of equation 8 exists.

[3] see FELLER[7]

[4] $\Pr(A \mid B)$ means the probability of event A assuming the event B

stands for the bottom of the layer and, consequently, $d \geq 1$ for its thickness. The nontrivial part of our concept is the transition probability, which has to express the scattering of light. Usually, the transition probability is a function of the state ℓ and / or the time t. In our case, this is wrong because of a travelling photon has a direction "up" or "down"[5] independent of ℓ or t. From this point of view, scattering has to be understood as changing the direction from "up" to "down" or vice-versa.

For that reason, let p be the probability that the photon changes its direction between t and $t-1$ because of scattering effects, hence the transition probabilities of our process are given by

$$\lambda_{t-1,\ell} \stackrel{\text{def}}{=} \begin{cases} p & \text{if } X_{t-2} = \ell - 2 \\ q & \text{if } X_{t-2} = \ell \end{cases} \tag{13}$$

and

$$\beta_{t-1,\ell+1} \stackrel{\text{def}}{=} \begin{cases} p & \text{if } X_{t-2} = \ell + 2 \\ q & \text{if } X_{t-2} = \ell, \end{cases} \tag{14}$$

where, as usual, the complement of p is denoted by

$$q \stackrel{\text{def}}{=} 1 - p. \tag{15}$$

Note that the events

$$X_{t-2} = \ell,\ X_{t-1} = \ell - 1,\ X_t = \ell$$

and

$$X_{t-2} = \ell,\ X_{t-1} = \ell + 1,\ X_t = \ell$$

indicate a change in direction.

Remark. Obviously, the transition probabilities (13) and (14) do not define a Markov chain. Therefore, the question arises why we call our process a random walk which is normally understood as a Markov Chain. The first reason is that with an additional parameter for the direction our process can be reformulated as Markov Process easily which will be the subject of a forthcoming paper Secondly, our argumentation concept in the next section is very well-known and characteristic for random walks.

Because of our Non-Markovian-transition probability the initial condition

$$P_{0,0} \ = \ 1$$

has to be extended to

$$P_{0,0} \ = \ 1 \quad \text{and} \quad P_{-1,-1} \ = \ 1 \tag{16}$$

or

$$P_{0,0} \ = \ 1 \quad \text{and} \quad P_{-1,1} \ = \ 1. \tag{17}$$

Both versions are possible and have their own right. The condition (16) describes the physical situation when the photon is getting into the layer. The second will be induced by recurrent events. In the rest of the paper probabilities assuming (17) will be characterized with a bar.

Finally, the probability that the layer material absorbs the photon during one time unit is denoted by p_a and its complement $1-p_a$ with q_a.

Remark. Clearly, the constants q_a and p correspond to K and S. For that reason, we call q_a the *absorbtion coefficient*.

For simplification FRESNEL-effects at the boundaries are ignored in a first approach but can be integrated in a second step analog to (9).

Now, the remaining question is the appearance of the reflectance R. Obviously, R is equivalent to the probability that the photon leaves again the colorant layer before it was absorbed in the material. But more interestingly, this probability can be expressed as a classical problem connected to random walks, namely the *first-passage time problem* which is illustrated next.

For the moment, we ignore absorption, hence we suppose $q_a = 1$. Let (16) be the initial condition. Then we consider the event:

$$0 \leq X_1,\ 0 \leq X_2, \ldots, 0 \leq X_{t-1},\ X_t = -1 \tag{18}$$

This means that the photon leaves the layer at time t or in random walk terminology, *the first visit to -1* takes place at the t-th step, see FELLER [7] for a detailled description. Let w_t be the probability of the event (18). Then we seek for the generating function

$$R(z) = \sum_{t=0}^{\infty} w_t \cdot z^t. \tag{19}$$

Proposition: *The reflectance R is immediately determined by $R(z)$.*

To see this, let $w_t(q_a)$ the probability w_t for an arbitrary absorption coefficient q_a, $0 \leq q_a \leq 1$. Then, the reflectance R is given by

$$R = w_0(q_a) + w_1(q_a) + w_2(q_a) + \cdots \quad . \tag{20}$$

The definition of q_a implies

$$w_t(q_a) \ = \ w_t \cdot q_a^t$$

which together with (20) leads to:

$$R \ = \ \sum_{t=0}^{\infty} w_t(q_a) \ = \ \sum_{t=0}^{\infty} w_t \cdot q_a^t \ = \ R(q_a) \tag{21}$$

Therefore, our next aim is the calculation of $R(z)$ which is the subject of section 3.

[5] and only these directions, analog to KUBELKA-MUNK

3. Evaluating the Model

First of all, our notation has to be refined in order to consider the influence of the layer thickness. The constant $d \geq 1$ was defined as largest level inside the colorant layer. Then w_t^d means w_t for a given d, where $d = \infty$ has the same interpretation as in section 1. Furthermore, as mentioned earlier the bar in $\bar{w}_t^d$ states briefly the initial condition (17)

$$P_{0,0} = 1 \quad \text{and} \quad P_{-1,1} = 1.$$

The corresponding generating functions and random variables X_t are characterized in the same way. Since a travelling photon starting at level 0 needs an even number of steps for coming back to 0, the level -1 is only reachable for odd t, hence

$$\bar{w}_t^\infty = 0 = w_t^\infty \quad \text{for } t \text{ even.} \tag{22}$$

Immediately from the definition we obtain

$$\bar{w}_1^\infty = p \quad \text{and} \quad w_1^\infty = q \tag{23}$$

which induces

$$w_1 = \frac{q}{p} \cdot \bar{w}_1^\infty. \tag{24}$$

Note that for $t \geq 3$ the probabilities $\bar{w}_t^\infty$ and w_t^∞ are connected by a different identity

$$w_t^\infty = \frac{p}{q} \cdot \bar{w}_t^\infty \tag{25}$$

which can be verified easily by induction on t. Next, we observe for odd $t \geq 3$

$$\bar{w}_t^\infty = q^2 \bar{w}_{t-2}^\infty + p(\bar{w}_3^\infty \bar{w}_{t-4}^\infty + \cdots + \bar{w}_{t-2}^\infty \bar{w}_1^\infty) \tag{26}$$

or equivalent with (15) and (23)

$$\bar{w}_t^\infty = (1-2p)\bar{w}_{t-2}^\infty + p(\bar{w}_1^\infty \bar{w}_{t-2}^\infty + \cdots + \bar{w}_{t-2}^\infty \bar{w}_1^\infty).$$

In order to show (26) the event leading to $\bar{w}_t^\infty$ is divided into a set of mutually exclusive events $A_2, A_4, A_6, \ldots$. Clearly, for $t \geq 3$ the photon goes to 1 in the first step. Therefore, there exists a smallest subscript i, $1 < i < t$, with $X_i = 0$. For every i the event A contains three blocks of trials.

Block 1. The photon goes to 1 in the first step.

Block 2. The photon needs exactly $i-1$ further trials to reestablish the initial situation.

Block 3. It takes exactly $t-1$ further trials to reach -1.

These three events depend on non-overlapping blocks of transitions with, by definition, fixed values of X_0, X_1, X_{i-1}, X_i and are therefore mutually independent. Every block forms a recurrent probability of type $\bar{w}_t$ or w_t, where the decision between $\bar{w}_t$ and w_t is given by the values of X_0, X_1, X_{i-1} and X_i. From the definition, the events in block 2 have the probabilities q for $i = 2$ and w_{i-1}^∞ otherwise. Obviously, the events in block 3 can be expressed as $\bar{w}_{t-i}^\infty$. So the probability of the simultaneous realization of all three events is given by the product

$$q^2 \cdot \bar{w}_{t-2}^\infty$$

for $i = 2$ and

$$q \cdot w_{i-1}^\infty \cdot \bar{w}_{t-i}^\infty \overset{(25)}{=} q \cdot \frac{p}{q} \cdot \bar{w}_{i-1}^\infty \cdot \bar{w}_{t-i}^\infty$$

otherwise. Summing up over all possible i completes (26).

Now we are able to determin $R(z)$. We set

$$\bar{w}_0^\infty = 0$$

for convenience and

$$\bar{w}_1^\infty = p$$

was given in (23). Next note that

$$\bar{w}_1^\infty \bar{w}_{t-2}^\infty + \cdots + \bar{w}_{t-2}^\infty \bar{w}_1^\infty$$

is the $(t-1)$-th coefficient from

$$(\bar{R}^\infty(z))^2.$$

Hence, we infer from (26)

$$\begin{aligned} &\bar{R}^\infty(z) - pz \\ &= \sum_{t=2}^{\infty} \bar{w}_t^\infty \cdot z^t \\ &= \sum_{t=2}^{\infty} [(1-2p)\bar{w}_{t-2}^\infty + p(\bar{w}_1^\infty \bar{w}_{t-2}^\infty + \cdots + \bar{w}_{t-2}^\infty \bar{w}_1^\infty)] z^t \\ &= (1-2p) z^2 \bar{R}^\infty(z) + pz(\bar{R}^\infty(z))^2. \end{aligned}$$

The physically reasonable solution of this quadratic equation is

$$\bar{R}^\infty(z) = \frac{1-(1-2p)z^2 - \sqrt{(1-(1-2p)z^2)^2 - 4p^2z^2}}{2pz}.$$

In a similar way we find for w_t^∞

$$w_t^\infty = p^2 w_{t-2}^\infty + q(w_3^\infty w_{t-4}^\infty + \cdots + w_{t-2}^\infty w_1^\infty)$$

and

$$R^\infty(z) = \frac{1+(1-2p)z^2 - \sqrt{(1+(1-2p)z^2)^2 - 4(1-p)^2 z^2}}{2(1-p)z}.$$

which is our analog to (6). This solution can be modified in the sense of SAUNDERSON. Let p_i and p_o be the special values of p modelling the FRESNEL-effects in connection with transmitting the level 0 and write f and F instead of w and R. Then the convolution equation changes to

$$f_1^\infty = 1-p_\mathrm{i}, \qquad f_3^\infty = p_\mathrm{i}(1-p)p_\mathrm{o}$$

and

$$f_t^\infty = p_\mathrm{i}p_\mathrm{o}w_{t-2}^\infty + (1-p_\mathrm{o})(f_3^\infty w_{t-4}^\infty + \cdots + f_{t-2}^\infty w_1^\infty)$$

implying

$$F^\infty(z) = \frac{(1-p_\mathrm{i})z - (1-p_\mathrm{o}-p_\mathrm{i})z^2R^\infty(z)}{1-(1-p_\mathrm{o})zR^\infty(z)}.$$

Remark. The foregoing argumentation is typical for the well-known CATALAN NUMBERS

$$c_n = \frac{1}{n+1}\binom{2n}{n}$$

satisfying

$$c_n = c_0c_{n-1} + \cdots + c_{n-1}c_0$$

for $n > 0$. Their generating function $C(z)$ is given by

$$C(z) = \frac{1-\sqrt{1-4z}}{2z}.$$

A very interesting survey in connection with lattice paths can be found in [8].

Finally, lets have a glance at bounded d (without SAUNDERSON-corrections. By induction on t we see $\bar{w}_t^d = 0 = w_t^d$ for even t and $\bar{w}_t^1 = q^{t-1}p$ otherwise implying

$$\begin{aligned} \bar{R}^1(z) &= pz + pq^2z^3 + pq^5z^6 + \cdots \\ &= pz(1 + (q^2z^2) + (q^2z^2)^2 + \cdots) \\ &= \frac{pz}{1-q^2z^2}. \end{aligned}$$

For $d \geq 2$ the middle part of the recurrent events have to be adapted to $d-1$, hence

$$\bar{w}_t^d = q^2\bar{w}_{t-2}^d + p(\bar{w}_3^{d-1}\bar{w}_{t-3}^d + \cdots + \bar{w}_{t-2}^{d-1}\bar{w}_1^d)$$

and

$$\begin{aligned} \bar{R}^d(z) &= \frac{pz}{1-(1-2p)z^2 - pz\bar{R}^{d-1}(z)} \\ &= \frac{1}{\underbrace{\frac{1-(1-2p)z^2}{pz}}_{\stackrel{\mathrm{def}}{=}u} - \bar{R}^{d-1}(z)} \\ &= \frac{1}{u-\bar{R}^{d-1}(z)}. \end{aligned}$$

This shows that for bounded d the generating functions $R(z)$ have the form of continued fractions, see for survey [9]. Because of their regular structure a closed form solution can be expected in near future.

4. Final Remarks

Obviously, the presented approach can be improved in many ways. Nevertheless, some interesting chances are offered, in particular, the generating function of first visit probabilities as a powerful tool to understand reflectance and the hope of an extended, higher dimensional light scattering theory.

References

[1] P. Kubelka and F. Munk. Ein Beitrag zur Optik der Farbanstriche. *Zeits. f. techn. Physik*, (12):593–601, 1931.

[2] F. Berg. *Isotrope Lichtstreuung in Papier - Neue Überlegungen zur Kubelka-Munk-Theorie*. PhD thesis, Technische Hochschule Darmstadt, April 1997.

[3] S. Mourad, P. Emmel, K. Simon, and R. Hersch. Extending Kubelka-Munk's Theory with Lateral Light Scattering. In D. Weiss and M. Yuasa, editors, **NIP 17***: International Conference on Digital Printing Technologies*, pages 469–473, Fort Lauderdale, Florida, September 30 – October 5, 2001. IS&T.

[4] W. Foote. An investigation of the optical scattering and absorption coefficients of dyed handsheets and the application of the I.C.I. system of color specification to these handsheets. *Paper Trade J.*, 108(TS125), March 9 1939.

[5] W. Foote. An investigation of the optical scattering and absorption coefficients of dyed handsheets. *Paper Trade J.*, 109(TS31), Dec. 21 1939.

[6] J. L. Saunderson. Calculation of the Color of Pigmented Plastics. *J. Opt. Soc. Amer.*, 32(4):727–736, 1942.

[7] W. Feller. *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons, 1968.

[8] R. Guy. Catwalks, Sandsteps and Pascal Pyramids. *Journal of Integer Sequences*, 3(Article 00.1.6), 2000.

[9] R. Graham, D. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison-Wesley, 1994.