

LETTER TO THE EDITOR

Analysis of airplane boarding via spacetime geometry and random matrix theory

E Bachmat¹, D Berend^{1,2}, L Sapir³, S Skiena⁴, N Stolyarov¹

¹ Department of Computer Science, Ben-Gurion University, Beer-Sheva 84105, Israel

² Department of Mathematics, Ben-Gurion University, Beer-Sheva 84105, Israel

³ Department of Management and Industrial Engineering, Ben-Gurion University, Beer-Sheva 84105, Israel

⁴ Department of Computer science, SUNY at Stony Brook, Stony Brook NY, 11794, USA

E-mail: ebachmat@cs.bgu.ac.il

PACS numbers: 02.10.Yn, 02.40.-k, 02.50.-r, 05.40.-a, 89.40.Dd

Abstract. We show that airplane boarding can be asymptotically modeled by two dimensional Lorentzian geometry. Boarding time is given by the maximal proper time among curves in the model. Discrepancies between the model and simulation results are closely related to random matrix theory. The models can be used to explain why some commonly practiced airline boarding policies are ineffective and even detrimental.

Submitted to: *J. Phys. A: Math. Gen.*

Airplane boarding is a process which is personally experienced daily by millions of passengers worldwide. Airlines have developed various strategies in the hope of shortening boarding time, typically leading to announcements of the form “passengers from rows 40 and above are now welcome to board the airplane”, often heard around airport terminals. We will show how the airplane boarding process can be asymptotically modeled by spacetime geometry. The discrepancies between the asymptotic analysis and finite population results will be shown to be related in some cases to random matrix theory (RMT). Previously, airplane boarding has been studied only empirically, via discrete event simulations [1, 2, 3].

We model the airplane boarding process as follows: Passengers $1, \dots, N$ are represented by coordinates $X_i = (q_i, r_i)$, where q_i is the index of the passenger along the boarding queue (first, second, third and so on), and r is his/her assigned row number. We re-scale (q, r) to $[0, 1] \times [0, 1]$. We assume that the main cause of delay in airplane boarding is the time it takes passengers to organize their luggage and seat themselves, once they have arrived at their assigned row. The input parameters for our model are:

u – the average amount of aisle length occupied by a passenger.

w – the distance between successive rows.

b – the number of passengers per row.

D – the amount of time (delay) it takes passengers to clear the aisle, once they have arrived at their designated row.

$p(q, r)$ – the joint distribution of a passenger’s row and queue joining time. This quantity is directly affected by the airline policy and the way passengers react to the policy.

For the purposes of presentation we shall assume that u, w, b, D are all fixed. The airplane boarding process produces a natural partial order relation of blocking among passengers. A passenger X *blocks* passenger Y if it is impossible for passenger Y to reach his/her assigned row before passenger X (and others blocked by X) has sat down and cleared the aisle. Airplane boarding functions as a peeling process for the partial order defined by the blocking relation. At first, passengers which are not blocked by any other passengers sit down. These passengers are the minimal elements in the blocking relation. In the second round, passengers which are not blocked by passengers other than those of the first round are seated and so forth. Boarding time thus coincides with the size of the longest chain in the partial order.

We assign to the boarding process with parameters $u, b, w, D, p(q, r)$ a Lorentz metric defined on the (q, r) unit square

$$ds^2 = 4D^2 p(q, r) (dqdr + k\alpha(q, r)dq^2), \quad (1)$$

where $k = bu/w$ and $\alpha(q, r) = \int_r^1 p(q, z)dz$. There are two properties of the metric which relate it to the boarding process:

(M1) The volume form of the metric is proportional to the passenger density distribution $p(q, r)$.

(M2) The blocking partial order among passengers during the boarding process

asymptotically coincides with the past-future causal relation induced by the metric on the passengers viewed as events in spacetime via their q, r coordinate representation.

To establish the second property, consider passengers represented by $X = (q, r)$ and $X' = (q + dq, r + dr)$, $dq > 0$. Consider the time when passenger X arrives at his/her designated row. All passengers with row numbers beyond r , which are behind passenger X in the queue but in front of passenger X' , will occupy aisle space behind passenger X . The number of such passengers is roughly $N\alpha dq$. Each such passenger occupies u/w units of aisle length, where we take the basic aisle length unit to be the distance between successive rows. The row difference between X and X' is $-(N/b)dr$. We conclude that passenger X is blocking passenger X' , via the passengers between them, roughly when $dq \geq -\alpha kdr$, a condition which coincides (together with $dq > 0$) with the causal relation induced by the metric.

By the two main properties, we may approximate asymptotically the airplane boarding process by the peeling process applied to the past-future causal relation on points in the associated spacetime, sampled with respect to the volume form. By a well-known result, 2-dimensional Lorentzian metrics are conformally flat. Hence, after an appropriate coordinate transformation, we may assume that the spacetime is given by a metric of the form

$$ds^2 = r(x, y)dxdy \tag{2}$$

on some domain (not necessarily the unit square). In the new coordinates x, y , which are light-like coordinates, chains in the causal relation coincide with increasing (upright) subsequences, namely, sequences of points (x_i, y_i) such that $x_i \leq x_j$ and $y_i \leq y_j$ for $i < j$. The peeling process applied to the causal relation coincides in this case with patience sorting, which is a well-known card game process, optimally computing the longest increasing subsequence in a permutation [4, 5].

Let $T(X)$ be the maximal proper time (integral over ds) of a time-like trajectory ending at $X = (q, r)$, and $L(\tau)$ be the length (integral over $\sqrt{-ds^2}$) of the space-like curve defined by the equation $T(X) = \tau$.

Using the analysis of the size of maximal increasing subsequences given in [7, 6], the two basic properties lead to the following modeling statement.

The boarding time of passenger i is approximately $\sqrt{N}T(X_i)$. In particular, the total boarding time is approximately $\sqrt{N} \max_X T(X)$, where $\max_X T(X)$ is the maximal proper time of a curve in the unit square with respect to the Lorentzian metric.

Here the word ‘‘approximately’’ means that the ratio tends to 1 with probability approaching 1 as the number of passengers N tends to infinity.

We apply the statement to the analysis of boarding times. Consider first the case where the airline does not have a boarding policy, namely passengers queue at uniformly

random times. Then $p(q, r) = 1$, $\alpha = 1 - r$, and the corresponding metric is

$$ds^2 = 4dq(dr + k(1 - r)dq) \quad (3)$$

We use this model to study the effect of airplane design parameters, such as the distance between rows and the number of passengers per row, on boarding time. These parameters affect boarding time through the parameter $k = \frac{ub}{w}$. Let $d = \max_X T(X)$ be the diameter of the model. Letting $A = (0, 0)$ and $B = (1, 1)$, we note that any point C in the unit square satisfies $A \prec C \prec B$, where \prec denotes the causal relation induced by (3). We conclude that the maximal curve must begin at A and end at B . To find the maximal proper time curve we solve the Euler-Lagrange equation for proper time, subject to the constraints of lying in the unit square and passing through A and B . Solutions to the unconstrained Euler-Lagrange (geodesic) equation are of the form

$$r = c_1 e^{2kq} + c_2 e^{kq} + 1. \quad (4)$$

For $k \leq \ln 2$ the geodesic solution passing through A and B is contained in the interior of the unit square, and is therefore the maximal curve. The length of the curve is

$$d(k) = 2\sqrt{\frac{e^k - 1}{k}}. \quad (5)$$

For $k > \ln 2$ the maximal curve contains a portion which runs along the q -axis until reaching the point $(1 - \ln 2/k, 0)$. From this point the curve follows a geodesic to the point $(1, 1)$. The geodesic is tangent to the q -axis at $(1 - \ln 2/k, 0)$, and this condition is in fact used to determine the point of departure from the boundary. We have

$$d(k) = 2\sqrt{k} + 2(1 - \ln 2)/\sqrt{k}. \quad (6)$$

We compared the above computations with simulations of our model of the boarding process. When $N = 10^6$, the length of the maximal chain matches well the estimated values derived from the space-time model. For $k = 0.5$, the length of the maximal chain in the simulation was 2261, while the estimated value is 2278; for $k = 5$, the corresponding values were 4589 and 4740, respectively.

We now consider more realistic values of N , namely $N = 100$ and $N = 200$. Table I presents the average boarding time results for 1000 computer simulations of the boarding process for several settings of k , compared with the boarding time estimate computed via the spacetime model. Upon inspection, Table I reveals that, for realistic values of N , there are differences in the range of 20-60 percent between the asymptotic boarding time estimates computed via Lorentzian geometry and the boarding time computed via simulations of the boarding process. We note the important fact that in all cases the Lorentzian estimate is larger.

To analyze the differences between our Lorentzian estimates and the simulation results, we consider L_N , the random variable representing the boarding time according to the boarding process model (the simulation results). We define the discrepancy random variable $M_N = L_N - \sqrt{N}d(k)$, which measures the difference between the boarding time and the Lorentzian estimate. The curvature of the metric given in (3) vanishes, and

k	N	simulation result	Space-time estimate
0.5	100	18.1	22.8
0.5	200	26.8	32.2
2.0	100	23.0	32.5
2.0	200	34.7	45.8
5.0	100	29.0	47.5
5.0	200	44.9	66.9

Table 1. A comparison of space-time model estimates with the average over 1000 simulations of boarding time results

therefore we can apply a coordinate transformation W which changes the metric to the form

$$ds^2 = 4dxdy. \quad (7)$$

Let S be the unit square and consider its image $U = W(S)$ under W . We refer to Figure 1 for an illustration of the following arguments. We recall that any point $C \in S$ satisfies $A \prec C \prec B$, where $A = (0, 0)$ and $B = (1, 1)$. We conclude that S will be mapped into the set of points D satisfying $W(A) \prec D \prec W(B)$, where now \prec denotes the causal relation with respect to (7). The set of such points is a rectangle R with sides parallel to the axis and corners at $W(A)$ and $W(B)$. W maps light-like curves with respect to (3) into light-like curves with respect to (7), which are simply vertical and horizontal lines. It also maps the geodesics of (3), which are given by (4), into straight lines. The left, right and top sides of S are light-like, and will therefore be mapped to the left, right and top sides of R . On the other hand, the bottom of S is not light-like nor a geodesic, and will therefore be mapped to a curve within R which will be concave since, by (4), the geodesics connecting points on the bottom of S lie below S . When $k \leq \ln 2$, the geodesic connecting A and B lies entirely within S . The image of this geodesic is the diagonal of R , and therefore $U = W(S)$ will contain the above diagonal triangle T within R . When $k > \ln 2$, the geodesic is not contained in S , and consequently the diagonal of R is not contained in U .

We can also compute the volumes of U and R as follows. Since W preserves volumes, the volume of U with respect to (7) equals the volume of S with respect to (3). The latter volume can be computed directly from the volume form formula $\sqrt{-\det(g)}dqdr$ and we get $\text{vol}(U) = 2$. The volume form formula also tells us that the volume with respect to (7) is twice the usual Euclidean area. Letting δx and δy denote the coordinate differences between $W(A)$ and $W(B)$, we have $\text{vol}(R) = 2\delta x\delta y$. The length of the diagonal of R is $\ell(\text{diag}) = 2\sqrt{\delta x\delta y}$, and therefore $\text{vol}(R) = \ell(\text{diag})^2/2$. On the other hand, the diagonal is simply the image under W of the geodesic connecting A and B , whose length by (5) is $2\sqrt{\frac{e^k-1}{k}}$. We conclude that $\text{vol}(R) = 2\frac{e^k-1}{k}$.

As noted previously, for the metric (7), the past-future causal chains correspond to increasing sequences of points. We also note that, since the spacetime points are sampled

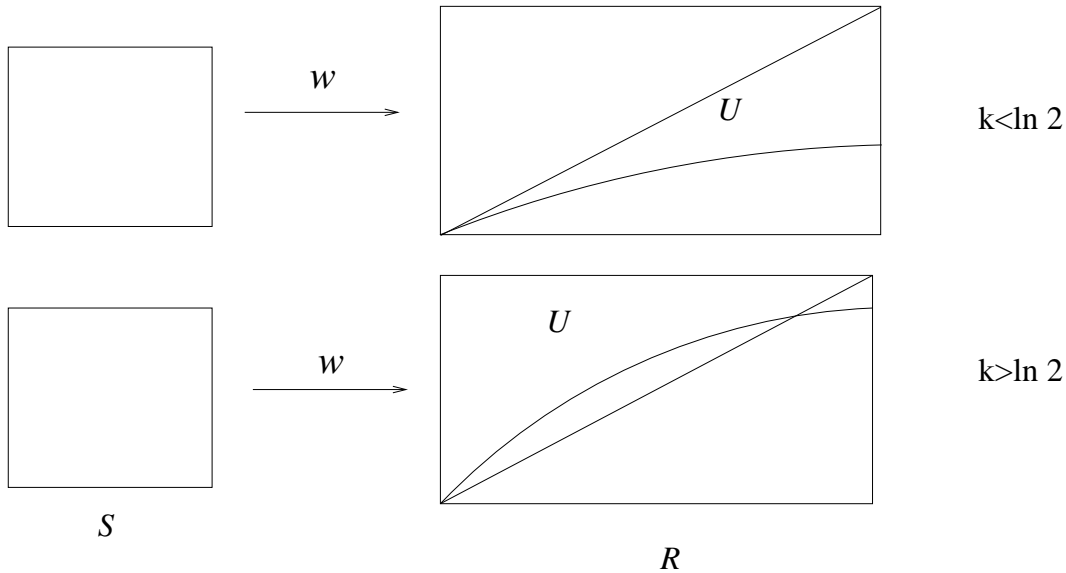


Figure 1. The image of the unit square under W .

according to the volume form, they are uniformly distributed. The problem is therefore reduced to that of computing discrepancies of increasing subsequences for N uniformly distributed points in $U = W(S)$. Discrepancies have been studied in the context of increasing subsequences of uniformly distributed points in the rectangle R [8], and in the above diagonal triangle T within R [9]. Following [8] it is convenient to consider the poissonized version, where the number of chosen points is given by a Poisson distribution with parameter N . In both cases the discrepancy M_N has order of magnitude $N^{1/6}$. We therefore define the scaled discrepancy $\Delta_N = M_N/N^{1/6}$. In the case of R , the scaled discrepancy is given asymptotically by the Tracy-Widom distribution F_2 [10], which measures the (scaled) discrepancy of the largest eigenvalue of an $N \times N$ matrix in the Gaussian unitary ensemble (GUE) [11], in comparison with $2\sqrt{N}$. For $N/2$ uniformly distributed points in T , the normalized discrepancy is given asymptotically by the Tracy-Widom Distribution F_4 [10], which is the normalized discrepancy of the largest eigenvalue in the Gaussian symplectic ensemble (GSE). The averages for these distributions are $E(F_2) \approx -1.77$ and $E(F_4) \approx -2.3$. After poissonization we may think of Poisson(N) uniformly distributed points in U as the restriction to U of Poisson($N \frac{\text{vol}(R)}{\text{vol}(U)}$) = Poisson($N \frac{e^k - 1}{k}$) uniformly distributed points in R .

The results of [8, 9] remain the same after depoissonization and can therefore be applied to the original version of our problem. When $k \leq \ln 2$ we have $T \subset U \subset R$, so that by applying the refined estimates to $N \frac{e^k - 1}{k}$ points in R we obtain

$$F_4(z) \leq Pr(\Delta_N \leq (\frac{e^k - 1}{k})^{1/6} z) \leq F_2(z). \quad (8)$$

In particular, we obtain the finer asymptotic estimate

$$-2.3(\frac{e^k - 1}{k})^{1/6} \leq E(\Delta_N) \leq -1.77(\frac{e^k - 1}{k})^{1/6}. \quad (9)$$

Applying the refined estimates to the results from table I with $k = 0.5$, we see that the refined estimate (9) holds already for the realistic values $N = 100, 200$. Indeed, $17.7 < 18.1 < 18.8$ for $N = 100$ and $26.4 < 26.8 < 27.8$ for $N = 200$.

The discrepancy M_N experiences a phase transition in behavior at the value $k = \ln 2$. For example, when $k = 3$, linear regression produced a tight fit with the formula

$$E(M_N) \sim -4.85N^{0.222}. \quad (10)$$

We do not know how to compute analytically the correct order of magnitude of M_N when $k > \ln 2$. However, we can provide the following heuristic explanation for the change in discrepancy behavior. When $k > \ln 2$, the maximal curve contains a portion of the bottom edge of the unit square. The argument essentially states that the form of the metric (3) near the bottom edge forces the portion of the maximal chain along the bottom edge to have height (transversal) fluctuations which are too small to produce a small discrepancy.

Consider the length of the longest chain between the point $A = (0, 0)$ and a second point $Z = (0, z)$ on the bottom edge of the unit square. The Lorentzian estimate for the length of the sequence is $2\sqrt{k}q\sqrt{N}$. We may consider the narrow band $B_a = \{(q, r) \mid 0 \leq r \leq N^{-a}\}$ on the bottom of the unit square. Assume that a given chain between A and Z has a non negligible portion which is outside B_a , namely, there are numbers $0 < c < d < z$ such that all points (q, r) in the chain with $c < q < d$ satisfy $r > N^{-a}$. On the complement of B_a consider the constant metric

$$ds^2 = 4dq(dr + k(1 - N^{-a})dq). \quad (11)$$

Any causal chain (in the complement of B_a) with respect to the metric (3) is also a chain with respect to (11). The longest chain with respect to (11) between (c, N^{-a}) and (d, N^{-a}) has, with high probability (w.h.p), a size of at most $2\sqrt{k}(c-d)\sqrt{N} - O(N^{1/2-a})$. The chain will therefore have w.h.p a negative discrepancy whose order of magnitude is at least $N^{1/2-a}$. We see that, whenever $a < 1/3$, a maximal chain which is not essentially confined to B_a will have w.h.p negative discrepancy whose order of magnitude is at least $N^{1/2-a} \gg N^{1/6}$. On the other hand, it follows from the computations in [12] that, whenever $a > 1/6$, a chain which is confined to B_a will produce w.h.p a negative discrepancy whose order of magnitude is larger than $N^{1/6}$. Putting these arguments together for $a = 1/4$, we obtain a strong heuristic evidence that the discrepancy is negative with order of magnitude larger than $N^{1/6}$.

The space-time metrics given by (1) can be used for comparing different boarding policies. Boarding policies affect the passenger distribution function $p(q, r)$. For example, assume that the policy is to board the back third of the airplane, followed by the front third and finally the middle third. The corresponding distribution $p(q, r)$ is given by $p(q, r) = 3$ on the sub-squares $[0, 1/3] \times [2/3, 1]$, $[1/3, 2/3] \times [0, 1/3]$, $[2/3, 1] \times [1/3, 2/3]$, and $p(q, r) = 0$ elsewhere. When $p(q, r) = 0$, the metric given in (1) degenerates. This problem can be overcome by setting $p(q, r) = \varepsilon$ in such regions and letting ε tend to 0. An equivalent and simpler approach is to allow the expression

$dqdr + k\alpha(q, r)dq^2$ from (1) to determine a light cone at all points. The airplane boarding time is then given asymptotically by the maximal proper time among all timelike curves even when $p(q, r)$ vanishes in certain regions.

A comparison of the results of the spacetime computations, with the parameter $k = 4$, and the results of detailed event driven simulations of boarding processes [2, 3], shows that the spacetime estimates are in almost complete agreement with the event driven simulation results regarding the ranking of the different policies. This is somewhat surprising given that the trace driven simulations in [2] take into account many details of actual boarding processes which are not considered by our model of the boarding process. The large discrepancies which we have noted previously between our own model of the boarding process and the spacetime estimates are less of a factor since, when comparing boarding strategies, only ratios of boarding times matter, and these are less affected by the discrepancies since in all cases the spacetime models lead to overestimates. Following the analysis we conclude that:

- (1) – The commonly practiced back-to-front boarding policies, which board passengers from the back of the airplane first, are ineffective for realistic values of $3 < k < 5$.
- (2) – Among row dependent policies which do not severely constrain passengers, random boarding (no policy) is almost optimal.

Finally, our methods can be applied to yield new insights into several other discrete processes including patience sorting [5], polynuclear growth models (PNG) [13], kinematics of causal sets [14], maximal layers in planar domains [15] and scheduling of I/O requests to disk drives [16, 17].

Acknowledgments

We are grateful to Perci Deift, Ofer Zeitouni and Jinho Baik for very useful discussions. This research has been partially supported by a grant from the dean of natural sciences at Ben-Gurion U. The first author has been partially supported by an IBM faculty award.

References

- [1] S. Marelli, G. Mattocks and R. Merry, *Boeing Aero Magazine* **1**, (2000).
- [2] H. Van Landegham and A. Beuselinck, *Euro. J. of Op. Res.* **142**, 294 (2002).
- [3] M. van den Briel, J. Villalobos, G. Hogg, T. Lindemann and A. Mule, *Interfaces* **35**, 191 (2005).
- [4] C.L. Mallows, *Bull. Inst. Math. Appl.* **9**, 216 (1973).
- [5] D. Aldous and P. Diaconis, *Bull. A. Math. Soc.* **36**, 413 (1999).
- [6] J.D. Deuschel and O. Zeitouni, *Ann. of Prob.* **23**, 852 (1995).
- [7] A. Vershik and S. Kerov, *Soviet Math. Dokl.* **18**, 527 (1977).
- [8] J. Baik, P. Deift and K. Johansson, *J. Am. Math. Soc.* **12**, 1119 (1999).
- [9] J. Baik and E. Rains, *Duke J. of Math.* **109**, 205 (2001).
- [10] C.A. Tracy and H. Widom, *Commun. Math. Phys.* **159**, 151 (1994); **177**, 727 (1996).
- [11] M.L. Mehta *Random Matrices*, (Academic press, 2004).
- [12] K. Johansson, *Commun. Math. Phys.* **209**, 437 (2000).
- [13] M. Prahofer and H. Spohn, *Phys. Rev. Lett.* **84**, 4882 (2000).

- [14] L. Bombelli, J. Lee, D. Meyer and R.D. Sorkin, *Phys Rev. Lett.* **59**, 521 (1987).
- [15] L. Devroye, *Computers and Math. with Applications* **25**, 19, (1993).
- [16] M. Andrews, M. Bender and L. Zhang, *Algorithmica* **32**, 277 (2002).
- [17] E. Bachmat, *Proc. of Symp. Th. of Comp. (STOC)*, 277 (2002).