Proceedings of the 48th European Study Group
Mathematics with Industry
Delft, 15-19 March 2004

Cor Kraaikamp, Hai Xiang Lin, Kees Oosterlee, editors

Contents

Preface

The forty-eighth study group Mathematics with Industry, which was held in Delft from 15-19 March 2004, was–as is usually the case with these study groups, ever since they started in the sixties at the University of Oxford–a huge success.

In Delft seven problems were addressed. The first three dealt with problems with a statistical nature. For instance, there was a problem from AOT, on option trading. The second problem was a problem posed by the Dutch Court of Audit (the 'Algemene Rekenkamer'), on how one should sample. This is an important problem, which seem to a interesting subject for further research. The third problem was supplied by the Dutch Bureau of Statistics (CBS), and dealt with the problem of statistical disclosure control. By law the CBS cannot disclose her sources–people like you and me–while she has to publish statistically relevant information for the Dutch society. The four other problems were of an analytic nature. They involved mathematical modeling of physical problems using (partial) differential equations, numerical methods, and theoretical analysis. Six problems are included in this proceedings, and the seventh problem was completely solved but not reported here.

All participants were very enthusiastic about this study week. This was apparent not only in the flurry of activities during the week at various floor of the 'EWI-building,' where the study week was held, but also in the quality of the presentations, or in the merry atmosphere during the diner of all the participants on Wednesday evening.

The results were presented by the various groups on the last Friday of the study week. Already during the week some of the subjects of the study week drew considerable attention. For example, the group leader Ludolf Meester of the audit problem was interviewed by the popular scientific magazine 'Natuur en Techniek,' on the findings of this group. Also the problem dealing with the 'leaning' of the Old Church in Delft drew a lot of media attention. Among others the 'Volkskrant', and the science magazine 'Noorderlicht' of the broadcasting-organization VPRO reported on this subject.

This study group could not have been organized without the help of the sponsors, and the participating companies supplying the problems, and often also man-power to tackle these problems. We, as organisors, are deeply indebted to them. In particular we would like to thank Filtrix and X-Flow, the Dutch Court of Audit, the CBS, AOT, Teijin Twaron, and Demis BV. Finally we would like to thank the department of archives of the municipality of Delft.

C. Kraaikamp, H.X. Lin, C.W. Oosterlee, editors and organizors.
Delft, November 22, 2005.

List of participants

Abramyan, Andrey	IPME RAS	andabr33@yahoo.co.uk
Almendral, Ariel	TU Delft	
Anderluh, Jasper	AOT	janderluh@aot.nl
Aspers, Wim	UL	wima@liacs.nl
Bremer, Marc	TU Delft	M.W.A.Bremer@ewi.tudelft.nl
Broeze, Ed	Alg. Rekenk.	e.broeze@rekenkamer.nl
Broomans, Peterjan	TU Delft	p.broomans@wbmt.tudelft.nl
Cator, Eric	TU Delft	e.a.cator@twi.tudelft.nl
Dekker, Kees	TU Delft	K.Dekker@ewi.tudelft.nl
Den Decker, Piet	Teijin Twaron	piet.dendecker@twaron.com
Fokkink, Robbert	TU Delft	r.j.fokkink@ewi.tudelft.nl
Grashoff, Poul	DEMIS	poulg@demis.nl
Hanzon, Bernhard	UL	bhanzon@planet.nl
Hendriks, Harrie	KUN	H.Hendriks@math.kun.nl
Hensbergen, Andr	TU Delft	A.T.Hensbergen@ewi.tudelft.nl
Hijmissen, Jelle	TU Delft	hijmissen@hotmail.com
Hof, Bas van 't	VORTech	bas@vortech.nl
Horssen, Wim van	TU Delft	w.t.vanhorssen@ewi.tudelft.nl
Jain, Sudhir	Aston U	S.Jain@aston.ac.uk
Janssen, Rik	STW	rik@stw.nl
Knoester, Henk	Teijin Twaron	henk.knoester@twaron.com
Koppelschaar, Carl	NOW	carl.koppeschaar@kennislink.nl
Kraaikamp, Cor	TU Delft	c.kraaikamp@ewi.tudelft.nl
Laan, vdr Pepijn		pepijnvanderlaan@planet.nl
Leentvaar, Coen	TU Delft	c.c.w.leentvaar@ewi.tudelft.nl
Lensink, Michiel	Filtrix	michiel.lensink@filtrix.com
Lin, Hai Xiang	TU Delft	h.x.lin@ewi.tudelft.nl
Malakpoor, K	TUE	k.malakpoor@tue.nl
Matusiak, Ewa	U Vienna	e-matusiak@yahoo.com
Meerman, Hans	Teijin Twaron	hans.meerman@twaron.com
Meester, Ludolf	TU Delft	L.E.Meester@ewi.tudelft.nl
Metrikine, Andrei	TU Delft	A.Metrikine@citg.tudelft.nl
Mohammadi, Leila	UL	leila@math.leidenuniv.nl
Mokveld, Philip	UvA	pmokveld@science.uva.nl
Naifar, Fahmi	TU Delft	f.naifar@ewi.tudelft.nl
Nuyens, Misja	UvA	mnuyens@science.uva.nl
Oosterlee, Kees	TU Delft	C.W.Oosterlee@ewi.tudelft.nl
Peletier, Mark	CWI	Mark.Peletier@cwi.nl
Perez, Etelvina Javierre	TU Delft	e.j.perez@ewi.tudelft.nl
Permana, Ferry	TU Delft	f.j.Permana@ewi.tudelft.nl
Perot, Blair	Umass	perot@ecs.umass.edu
Pik, Derk	UL	drpik@math.leidenuniv.nl

Planque, Bob	CWI	rplanque@cwi.nl
Poulisse, Hennie	Shell	Hennie.Poulisse@shell.com
Prokert, Georg	TUE	g.prokert@tue.nl
Rottschfer, Vivi	UL	vivi@math.leidenuniv.nl
Rozenholc, Yves	U Paris	yves.rozenholc@math.jussieu.fr
Savcenco, Valeriu	CWI	V.Savcenco@cwi.nl
Sonneveld, Peter	TU Delft	p.sonneveld@ewi.tudelft.nl
Verleye, Bart	KU Leuven	bart.verleye@cs.kuleuven.ac.be
Vermolen, Fred	TU Delft	F.J.Vermolen@ewi.tudelft.nl
Vuik, Kees	TU Delft	c.vuik@ewi.tudelft.nl
Wawrzyniak, Marta	TU Delft	m.wawrzyniak@ewi.tudelft.nl
Weide, Hans vd	TU Delft	j.a.m.vanderweide@ewi.tudelft.nl
Wesseling, Piet	TU Delft	p.wesseling@ewi.tudelft.nl
Williams, JF	CWI	williams@cwi.nl
Wilson, Charles	TU Delft	C.W.Mahera@ewi.tudelft.nl
Winden, Koos van	VU	koos@cybercom.net
Wolf, Peter-Paul de	CBS	pwof@cbs.nl
Zielman, Berrie	Alg. Rekenk.	A.Zielman@rekenkamer.nl

ADR Option Trading

Jasper Anderluh and Hans van der Weide

TU Delft, EWI (DIAM), Mekelweg 4, 2628 CD Delft
j.h.m.anderluh@ewi.tudelft.nl, J.A.M.vanderWeide@ewi.tudelft.nl

1. Introduction

A company that is seeking to raise capital to finance necessary investments, can issue stocks, which are basically certificates of partial ownership in the company. There are many rules for issuing stocks, one of which is that the company has its seat in the country where the stocks are issued. If, nevertheless, a non-US company, like Royal Dutch N.V., wants to raise capital in the US, it can issue ADRs. ADR is an acronym for American Depository Receipt, which is a certificate issued by a US bank, representing a certain amount of stock of a non-US company on a non-US exchange. Just as US stock, ADRs can be traded, cleared and settled on American exchanges in accordance with US market regulations. ADRs are US securities and are quoted and traded in US dollars. This makes it easier for Americans to invest in non-US companies, due to the widespread availability of dollar-denominated price information, lower transaction costs, and timely dividend distributions. The price of an ADR follows, accounting for the currency exchange rate, more or less the price in the home country; if the US price gets too far off from the price in the home country, arbitrageurs will step in the market and the arbitrage opportunity will soon cease to exist. In order to provide the American investor with more investment possibilities, options are issued on these ADRs. These ADR options are also listed on US markets, denominated in US dollars and also the strike is specified in US dollars.
Non-US market makers trading options on a stock listed in their domestic country might be interested in adding the corresponding ADR options to their portfolio. The interesting part of ADR option trading, is the integration of the position in these US listed options with the domestic option position. The advantage of this integration is that we have - from a risk-management point of view - a clear perspective of the exposure the market maker has with respect to a single stock. If we consider for example stocks Royal Dutch (RD), traded in Amsterdam and their corresponding US ADRs, we can - once we are able to manage this as one integrated position - compute a single delta, gamma or vega for our Royal Dutch position. Furthermore we would like to exploit the

mis-pricing of US options with respect to their Dutch counterparts and so we need a pricing model to price the foreign US dollar denominated ADR options consistently with the domestic Euro denominated options and stock.

The market model. We start building our market model from the domestic stock price process $\{S_t\}_{t\geq 0}$, that we model as a Geometric Brownian Motion,

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \quad S_0 = s_0 \quad \Leftrightarrow \quad S_t = s_0 e^{(\mu - 0.5\sigma^2)t + \sigma W_t}. \tag{1}$$

This is the classical approach to stock price modeling as is also used by [**2**]. For the Euro/Dollar exchange rate process $\{FX_t\}_{t\geq 0}$ we also assume that it is given by a GBM,

$$\begin{aligned} \frac{dFX_t}{FX_t} = \alpha dt + \Sigma_1 dW_t + \Sigma_2 dZ_t \quad FX_0 = f_0 \quad \Leftrightarrow \\ FX_t = f_0 e^{(r - 0.5(\Sigma_1+\Sigma_2)^2)t + \Sigma_1 W_t + \Sigma_2 Z_t}. \end{aligned} \tag{2}$$

Here we used another standard Brownian Motion Z independent of W to model a dependence structure between the domestic asset S and the exchange rate FX. This is the same approach as in [**3**] and [**5**]. We remark that the direction of the exchange rate is such that the value FX_t is the number of Euros you have to pay for one US dollar at time t. Denote the ADR stock price process by $\{A_t\}_{t\geq 0}$. We assume that the market is efficient, i.e. arbitrageurs are active to force the following relation to hold,

$$A_t = \frac{S_t}{FX_t} \quad t \geq 0. \tag{3}$$

This relation is investigated in [**1**] and turned out to be quite accurate looking at real markets where prices are formed concerning transaction and conversion costs. If we consider a European call option with strike K written on the ADR and therefore listed on the foreign market, the pay-off in US dollars Φ_C of this contract can be written using the previous relation by

$$\Phi_C(A_T) = \Phi_C\left(\frac{S_T}{FX_T}\right) = \left(\frac{S_T}{FX_T} - K\right)^+. \tag{4}$$

We remark that also the strike K is denominated in US dollars. Now we need to find the equivalent martingale measure $\mathbb{Q}$ turning all the assets in our economy into martingales in order to price this derivative. First we have to identify the assets we can use building our portfolio. As in the classical approach we use both the domestic stock S and

the domestic bank-account process $\{B_t^{(d)}\}_{t\geq 0}$ as assets in our economy, where $B^{(d)}$ is given by,

$$B_t^{(d)} = e^{r_d t}.$$

Here r_d is the classical risk-free rate. As an extra asset we introduce the foreign or US bank-account process $\{B_t^{(f)}\}_{t\geq 0}$. We are considering our economy from the domestic perspective (Euro-zone), so we should denominate all our assets in the same domestic currency and therefore we consider the US bank-account denominated in Euros as a risky asset. So the US bank account is not a bank account in the classical sense, i.e. from the domestic point of view it is not offering the risk-free rate. To illustrate this we take a closer look at trading of this asset, which is converting one Euro at time $t = 0$ into $(FX_0)^{-1}$ US dollars and deposit this amount on a US bank account. At time t we earned the US risk-free rate on the deposited dollar amount and in order to calculate its value in Euros we have to convert it again by the stochastic exchange rate FX_t. We have for $B^{(f)}$,

$$B_t^{(f)} = \frac{FX_t}{FX_0} e^{r_f t}.$$

Here r_f is the foreign risk-free rate, that we cannot obtain risk-free if we denominate the value in Euros. Using the dynamics of the exchange rate FX_t we can compute the dynamics of $B_t^{(f)}$ by,

$$\begin{aligned} dB_t^{(f)} &= FX_t e^{r_f t} r_f dt + e^{r_f t} dFX_t \\ &= B_t^{(f)} \left[(r_f + \alpha) dt + \Sigma_1 dW_t + \Sigma_2 dZ_t \right] \end{aligned}$$

From (3) we recognize that the ADR price process is completely determined by the domestic stock price process S and the exchange rate FX and therefore we do not want to introduce A as an extra asset in our economy. If we decide to choose $B^{(d)}$ as the numéraire, which is more or less a standard choice, we can find the option price by identifying the equivalent martingale measure $\mathbb{Q}$ such that the discounted asset price processes $S_t[B_t^{(d)}]^{-1}$ and $B_t^{(f)}[B_t^{(d)}]^{-1}$ are martingales. If we denote the discounted stock price process by $\tilde{S}$, we are looking for a measure $\mathbb{Q}$ such that both the process $\tilde{W}$ defined by,

$$\tilde{W}_t = W_t + q_1 t \quad \Leftrightarrow \quad d\tilde{W}_t = dW_t + q_1 dt$$

is a standard Brownian Motion and the discounted stock price process $\tilde{S}$ is a martingale. The existence of such a $\mathbb{Q}$ is guaranteed by the Girsanov Theorem, see e.g. [**4**]. Writing the dynamics of $\tilde{S}$ in terms of $\tilde{W}$ we get,

$$d\tilde{S}_t = \tilde{S}_t \left[(\mu - r_d) dt + \sigma dW_t \right] = \tilde{S}_t \left[(\mu - r_d - \sigma q_1)\, dt + \sigma d\tilde{W}_t \right].$$

For $\tilde{S}$ to be a martingale, we set the drift term equal to zero, so

$$q_1 = \frac{\mu - r_d}{\sigma}.$$

Now we have solved q_1 we directly obtain the dynamics of S under $\mathbb{Q}$ by,

$$\begin{aligned} dS_t &= S_t \left[\mu dt + \sigma dW_t\right] \\ &= S_t \left[\mu dt + \sigma d(\tilde{W}_t - q_1 dt)\right] = S_t \left[r_d dt + \sigma dW_t\right]. \end{aligned}$$

This is not a surprising result, because it is equivalent to the classical risk-neutral Black-Scholes dynamics of the stock price process, see [**2**]. Now we proceed by changing the drift of the other Brownian Motion Z such that both the process $\tilde{Z}$ defined by,

$$\tilde{Z}_t = Zt + q_2 t \quad \Leftrightarrow \quad d\tilde{Z}_t = dZ_t + q_2 dt$$

is a $\mathbb{Q}$ standard Brownian Motion, independent of $\tilde{W}$ and the process $\tilde{B}$ is a martingale. Here $\tilde{B}$ denotes the discounted foreign bank account process $B_t^{(f)}[B_t^{(d)}]^{-1}$. For the dynamics of $\tilde{B}$ we obtain,

$$\begin{aligned} d\tilde{B}_t &= \frac{1}{B_t^{(d)}} dB_t^{(f)} - \frac{\tilde{B}_t}{B_t^{(d)}} dB_t^{(d)} \\ &= \tilde{B}_t \left[(r_f - r_d + \alpha)dt + \Sigma_1 dW_t + \Sigma_2 dZ_t\right] \\ &= \tilde{B}_t \left[(r_f - r_d + \alpha)dt + \Sigma_1 (d\tilde{W}_t - q_1 dt) + \Sigma_2 (d\tilde{Z}_t - q_2 dt)\right] \\ &= \tilde{B}_t \left[\left(r_f - r_d + \alpha - \Sigma_1 \frac{\mu - r_d}{\sigma} - \Sigma_2 q_2\right) dt + \Sigma_1 d\tilde{W}_t + \Sigma_2 d\tilde{Z}_t\right]. \end{aligned}$$

Again we need the drift term equal to zero for $\tilde{B}$ a $\mathbb{Q}$ martingale, which is satisfied if we put,

$$q_2 = \frac{r_f - r_d + \alpha - \Sigma_1 \frac{\mu - r_d}{\sigma}}{\Sigma_2}.$$

Now we find for the exchange rate process FX under the pricing measure $\mathbb{Q}$:

$$\begin{aligned} dFX_t &= FX_t \left[\alpha dt + \Sigma_1 dW_t + \Sigma_2 dZ_t\right] \\ &= FX_t \left[(\alpha - \Sigma_1 q_1 - \Sigma_2 q_2)dt + \Sigma_1 d\tilde{W}_t + \Sigma_2 d\tilde{Z}_t\right] \\ &= FX_t \left[(r_f - r_d)\, dt + \Sigma_1 d\tilde{W}_t + \Sigma_2 d\tilde{Z}_t\right]. \end{aligned}$$

It is clear that the exchange rate process FX does not enter into our portfolio, because it is not possible to actually buy or sell the exchange rate. We can however keep an amount of foreign currency on a foreign bank account, that is why we decided to take the foreign bank account as a possible asset for our portfolio. We used the fact that all discounted

assets - where we used the domestic bank account as numéraire process - have to be martingales under $\mathbb{Q}$ to derive the dynamics of the exchange rate process FX under $\mathbb{Q}$. We can now use this measure and the corresponding asset-dynamics to price all attainable claims, expressed in the domestic currency. Suppose $\{X_t\}_{t\geq 0}$ is the vector containing all our assets in the economy, a claim Φ is attainable if we can find a self-financing strategy ϕ such that,

$$\phi_0 \cdot X_0 + \int_0^T \phi_u dX_u = \phi_T \cdot X_T = \Phi(X_T). \tag{5}$$

Here the self-financing property is used in the first equality sign. We used in our setup the domestic bank account as the numéraire and constructed the corresponding martingale measure. If we consider the discounted asset price process $\tilde{X}$ given by,

$$\tilde{X}_t = \frac{X_t}{B_t^{(d)}},$$

we have that the self-financing property of ϕ also holds for $\tilde{X}$. Now we use that the discounted asset prices are martingales and therefore - assuming the right conditions on ϕ - we have that the stochastic integral representing the gains of the strategy ϕ over time is also a martingale. We remark that we have to impose some conditions on ϕ to guarantee the stochastic integral to be a martingale instead of being a local martingale only. From a no-arbitrage argument we have that $\phi_0 \cdot X_0$ must be equal to the price of the contract at time 0. Now we can compute this price V_Φ of the option with pay-off Φ using the martingale property of the discounted asset prices by,

$$\begin{aligned} V_\Phi &= B_0^{(d)} \phi_0 \cdot \tilde{X}_0 = B_0^{(d)} \mathbb{E}_\mathbb{Q}\left[\phi_T \cdot \tilde{X}_T\right] \\ &= \frac{B_0^{(d)}}{B_T^{(d)}} \mathbb{E}_\mathbb{Q}\left[\Phi(X_T)\right] = e^{-r_d T} \mathbb{E}_\mathbb{Q}\left[\Phi(X_T)\right]. \end{aligned} \tag{6}$$

All the assets in our economy are in Euro currency and as we are using this assets to replicate our foreign option (4), we also have to denote the pay-off of that option in Euros $\tilde{\Phi}_C$ by,

$$\tilde{\Phi}_C(A_T) = FX_T \Phi_C(A_T) = \left(S_T - FX_T K\right)^+. \tag{7}$$

Now we use the pricing formula (6) to obtain $\tilde{C}_{for}$ the price of the foreign option in Euros:

$$\tilde{C}_{for} = e^{-r_d T} \mathbb{E}_\mathbb{Q}\left[FX_T \left(A_T - K\right)^+\right] = e^{-r_d T} \mathbb{E}_\mathbb{Q}\left[\left(S_T - FX_T K\right)^+\right]. \tag{8}$$

As this option is traded in the US market we need an option value in US dollars in order to compare our theoretical price to the market price. This is nothing else then converting the Euro value $\tilde{C}_{for}$ into a

value C_{for} in US dollars by dividing by FX_0. In this section we are considering European call options, whereas in practice the options on ADRS are American. A related matter of practical concern is the ex-dividend date of an ADR that is typically a few days earlier or later than the ex-dividend date of the non-US stock the ADR is based on.

Calibration of the Model to the market. If we want to use the model in practice, we should come up with a method to determine the parameters σ, Σ_1 and Σ_2 as they appear in (1) and (2). We will relate these parameters to the volatility of both the stock price process and the exchange rate process, where the volatility is defined as the standard deviation of the log-returns, scaled by the square-root of time up to time units of 1 year. Furthermore we relate the quantities we have to estimate to the correlation between the log-returns of the stock price process and the log-returns of the exchange rate process. The reason for doing so, is that we can obtain the volatilities and correlation as described above directly from an information system (e.g. Bloomberg) as these systems are used in a trading firm.
Suppose we have observations of the price processes at time points $\{t_0, t_1, .., t_M\}$, where we denote the fixed time step $t_{i+1} - t_i$ by Δt. Now we find for the log-return s_i at time t_{i+1} of the stock price process S,

$$s_i = \ln \frac{S_{t_{i+1}}}{S_{t_i}} = (\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma(W_{t_{i+1}} - W_{t_i}) \tag{9}$$
$$\stackrel{d}{=} (\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}U,$$

where $U \sim \mathcal{N}(0,1)$. If we denote the volatility estimate we obtain from our information system by $\hat{\sigma}^S$, we immediately can use it as an estimate for our σ. As we choose our domestic stock model as the standard GBM we of course expect that the volatility estimate from the information system is an estimator for our volatility parameter. The more interesting case is the determination of the parameters Σ_1 and Σ_2. For the exchange rate process FX we have a similar expression for the log-returns f_i,

$$f_i = \ln \frac{FX_{t_{i+1}}}{FX_{t_i}} = (\alpha - \frac{1}{2}(\Sigma_1 + \Sigma_2)^2)\Delta t + \Sigma_1(W_{t_{i+1}} - W_{t_i}) + \Sigma_2(Z_{t_{i+1}} - Z_{t_i})$$
$$\stackrel{d}{=} (\alpha - \frac{1}{2}(\Sigma_1 + \Sigma_2)^2)\Delta t + \sqrt{\Delta t}\,(\Sigma_1 U + \Sigma_2 V)\,. \tag{10}$$

Here we have again $V \sim \mathcal{N}(0,1)$, where U and V are independent. If we denote the estimate for the volatility of the exchange rate process by $\hat{\sigma}^{FX}$, we can write,

$$\hat{\sigma}^{FX} = \sqrt{\hat{\Sigma_1^2 + \Sigma_2^2}}. \tag{11}$$

We need more information to determine Σ_1 and Σ_2 separately. For the covariance between the log returns of the stock price process S and the exchange rate process FX at time t_{i+1} we have,

$$\mathrm{COV}(s_i, f_i) = \sigma \Delta t \mathrm{COV}(U, \Sigma_1 U + \Sigma_2 V) = \sigma \Delta t \Sigma_1.$$

For the correlation ρ we have

$$\rho = \frac{COV(s_i, f_i)}{\sigma \Delta t \sqrt{\Sigma_1^2 + \Sigma_2^2}} = \frac{\Sigma_1}{\sqrt{\Sigma_1^2 + \Sigma_2^2}}. \tag{12}$$

So using these relations we can, provided with estimates $\hat{\sigma^S}$, $\sigma^{\hat{F}X}$ and $\hat{\rho}$ from the information system, come up with estimates for the parameters in our model,

$$\sigma = \widehat{\sigma^S}, \quad \Sigma_1 = \hat{\rho}\widehat{\sigma^{FX}}, \quad \Sigma_2 = \widehat{\sigma^{FX}}\sqrt{(1-\hat{\rho})^2}. \tag{13}$$

Conclusion. In this paper we treat the topic of pricing options on ADRs listed on US markets. The importance is in the fact that the ADR price process is directly related to the price process of the corresponding stock in the non-US market. This relation extends to the possibility to replicate the foreign listed option with instruments in the domestic market. We set up an economy consisting of the domestic stock and bank account and moreover we introduced the foreign bank account as an additional risky asset. After modeling these instruments, we derived an equation for the price of a foreign call option. Finally we showed how we can relate standard estimates obtained from a trading information system to the parameters of our model.

References

[1] D.P. Miller, M.R. Morey *The Intraday Pricing Behavior of International Dually Listed Securities*, Journal of International Financial Markets, Institutions & Money, **6**(4), 79-89, (1996)

[2] F. Black, M. Scholes *The pricing of options and corporate liabilities*, J. Political Econ. **81**, 637-654, (1973)

[3] T. Björk, *Arbitrage Theory in Continuous Time*, 1. Edition, Oxford University Press, Oxford, (1998)

[4] I. Karatzas, S. E. Shreve, *Brownian Motion and Stochastic Calculus*, 2. Edition, Springer-Verlag, Berlin, (1991)

[5] M. Musiela, M. Rutkowski, *Martingale Methods in Financial Modelling*, 2. Printing, Springer-Verlag, Berlin, (1998)

Isolating and correcting errors while auditing accounts

Harrie Hendriks†, Cor Kraaikamp, Ludolf Meester‡, Philip Mokveld♮, Misja Nuyens♢

† Radboud Universiteit Nijmegen, P.O.Box 9010 6500 GL Nijmegen
hhendr@math.kun.nl
‡ TU Delft, EWI (DIAM), Mekelweg 4, 2628 CD Delft
c.kraaikamp@ewi.tudelft.nl, l.e.meester@ewi.tudelft.nl
♮ UvA, KdV-Instituut, Plantage Muidergracht 24, 1018 TV Amsterdam,
pmokveld@science.uva.nl
♢ VU, Department of Mathematics, De Boelelaan 1081, 1081 HV Amsterdam
mnuyens@few.vu.nl

Abstract. A practice in auditing of isolating and correcting errors as part of a statistical audit is introduced with an example. A brief literature review is presented. Decision procedures based on (Poisson) upper confidence bounds are analyzed, using the all-or-nothing model. Three situations are analyzed in detail, with primary focus on the probability that an approved account (still) contains a material error. This probability should not exceed the specified level (0.05 in this paper). It is found that in all these three cases it does, and, for some parameter values, substantially.

1. Introduction and problem statement

One of the problems presented to the 48th European Study Group *Mathematics with Industry* was posed by employees of the Netherlands Court of Audit (de Algemene Rekenkamer) and pertains to the auditing practice of isolating and correcting errors while auditing accounts. A long-standing dispute on the admissibility of some or all of these practices exists between the court and several departmental audit departments (departementale accountantsdiensten). The authors spent the week working on this problem; their findings are reported here. The limited time implies that not all angles could be covered, and many issues that were raised await further study, in particular whether the model adequately reflects common practice.

We start with an example. In a monetary unit sample, an auditor samples 100 Euros from an account of € 1 million. The *interval* J is the number of Euros in the account represented by one Euro in the sample, in this case $J =$ € 1 million$/100 =$ € 10 000. The *materiality*—setting an upper bound to the total error amount that is still acceptable—is set at 4% or € 40 000. As it happens, one error is found, in a 'personnel'

item. Common (statistical) procedure applied here is to compute an upper confidence bound for the total error, based on the Poisson distribution. With one error this bound would be 4.75 times the interval J, or €47 500. As this upper bound exceeds the materiality, approval would have to be withheld.
It happens to be the case, however, that this error could only occur in a 'personnel' item, and not in other items. Let us say, these are all 'materials' items. The auditor may decide to extrapolate (or project) this error only to the corresponding segment of the account; we shall employ the term *stratum* for such a segment. Furthermore, he may attempt to correct the error.
So, in hindsight, the account and the sample are viewed as follows:

Stratum:	Materials	Personnel
Total amount:	800 000	200 000
Number sampled:	80	20

In the materials stratum no errors are encountered, the corresponding Poisson upper bound therefore is $3J$ = €30 000. In the personnel stratum one error of size €5000 is found among the sampled items and it cannot be corrected. An exhaustive examination of the whole stratum follows, but no other errors are found, whence the total error in this stratum is known to be €5000.
The auditor may now combine the results in an overall upper bound €30,000+€5,000 = €35,000 that is below the materiality (€40,000). Based on this, the account could be approved. In case the error in the personnel item *could* be corrected, the adjusted bound would be €30,000, with the same result.

The question we consider is: *Is it permissible to modify regular statistical procedures with isolation and correction steps? That is, do these procedures retain their nominal properties such as confidence level and significance level?*

2. Isolating and correcting errors

For the sake of clarity we summarize what we mean by 'isolating and correcting errors' in this paper. In reality the range of these practices may be wider, which may cause our conclusions to be conservative. We believe that the following is in agreement with *International Standard on Auditing* 530 [**5**].
What does it mean to *isolate an error*? It should be noted that some errors can only occur in certain strata: the error in the example above could only occur in a personnel item. Isolating an error means the identification of a stratum for which the error is typical and extrapolate the error only to this stratum, rather than to the whole account.

A crucial issue here is whether strata and their boundaries are identified before or after the sample is drawn. In the case after, it may be possible to find a suitable stratum for every error found, and with some creative reasoning perhaps a suitably *small* stratum. After all, extrapolating an error to a small stratum leads to a smaller overall error estimate than when extrapolating it to the whole account.
At this point a modelling difficulty is encountered: how should one model an auditor who isolates and corrects 'in good faith' (so not 'too creatively')? We have tried to resolve this issue by exploring several situations, including two extremes: in Section 6 the strata are identified before the sampling and in Section 7 we assume that every error is isolated to a very small stratum.
Isolating errors may be combined with correcting them. It is natural to attempt to correct errors that are encountered and check for similar errors in similar book items and correct them as well. In the examples, we shall assume that the isolating stratum, when it is identified, is examined exhaustively and the errors found are corrected as much as possible. As a result, the exact error in this part of the account is known. While it is clear that these steps reduce the overall error amount, it is less clear how previously computed confidence bounds should be adjusted, or what these adjusted bounds mean.

3. A brief review of (some of) the literature

There are two papers—[**2**] and [**6**]—that discuss the theoretical aspects of isolating and projecting errors, and that report on the behavior of auditors when confronted with 'unique' errors, i.e., errors that are considered atypical for the population where the sample is taken from. Although these two papers—and also the papers [**3**] and [**4**]—report similar behavior of auditors ("a large majority of auditors favored isolating errors that appeared to be unique" ([**2**], p. 246)), the theoretical parts of these two papers express strongly opposing views.
The oldest of these two papers, Burgstahler and Jiambalvo [**2**], discusses a simple balls-in-urn example to show that "to 'isolate' some sample items found to be in error where 'isolated' items are not projected to the population [...] may lead auditors to systematically underestimate population error and result in excessive probability of incorrect acceptance" ([**2**], p. 234). In their balls-in-urn example Burgstahler and Jiambalvo only substantiate the first remark. In fact, their point of view is in part philosophical. In their view, "a fundamental assumption underlying audit sampling is that items are, for sampling purposes, homogeneous in the sense that observation of some subset of items is useful for drawing conclusions about the remainder of the whole population. [...] Further, this assumption is necessary for both statistical

and judgemental sampling in auditing." To isolate errors, and not projecting them seems to be in contradiction with this basic assumption of statistical auditing.

Their example is as follows. An urn contains black and red balls; the red balls correspond to items containing an error. The auditor is to construct the maximum likelihood estimate of the number of red balls in the urn. The maximum likelihood estimate is used in [**2**] because it is more intuitive than the more common confidence interval. Suppose there are N balls in the urn, of which an unknown number R are red, and the remaining $N-R$ are black. The number of red balls in a sample of size n drawn with replacement from this population of N balls has a binomial distribution with parameters n and R/N. The maximum likelihood estimate of R/N is r/n, where r is the number of red balls in the sample. This estimator is unbiased, an observation not made in [**2**]. Now suppose that the auditor draws r red balls, $n-r-1$ black balls, and one red cube (a 'unique' error). So one of the sample items is qualitatively different from the others. According to Burgstahler and Jiambalvo there are (at least) two approaches that might be adopted.

In the first approach one simply estimates the number of red items in the urn. This approach treats the red cube the same as a red ball; it is not isolated. The estimated number of red items in the population is $N(r+1)/n$.

In the second approach the red cube is assumed to be unique, and is not taken as an indication that other red items (such as cones, discs, etc.) might be in the urn. In this case it is assumed that of the items in the urn, one is known to be a red cube, an unknown number R of the balls are red, and the remaining $N-R-1$ items are assumed to be black balls. So the sample proportion of red balls is $r/(n-1)$, and the estimate of the total number of red items is $1+(N-1)r/(n-1)$.

Which is the correct approach? According to Burgstahler and Jiambalvo the problem with the second approach is that "auditors are seldom faced with situations where it is reasonable to rule out the possibility that a population contains other unique red shapes. [...] In auditing, no two errors are truly identical; each error is associated with *some* unique characteristic." ([**2**], p. 236).

According to Wheeler *et al.* ([**6**]), Burgstahler and Jiambalvo "suggested that an estimator in which no sample errors are isolated from the estimator (project all sample errors) is normative per standards of statistical inference" ... and "that an estimator biased by the isolation of nonrecurring errors violates those standards." ([**6**], p. 263). In the view of Wheeler *et al.* "a focus on bias ignores the dispersion of estimates." It is shown in [**6**], in an balls-in-urn example that in simplicity resembles that of Burgstahler and Jiambalvo, that in a situation

where there are very few red objects other than red balls, the biased estimator giving an estimate of the proportion of red balls in the population might be more desirable than the (unbiased) estimator giving an estimate of the number of red items in the population. It is shown that under certain circumstances the biased estimator has a smaller *mean squared error* than the unbiased estimator (the mean squared error equals the square of the bias plus the variance of the estimator). "A biased estimator may be more precise than an unbiased one" and evidence is provided that "the exclusion of unique errors in developing a sample estimator can increase the accuracy of the estimation process" ([**6**], p. 273).
Apparently, Wheeler *et al.* think to have refuted Burgstahler and Jiambalvo. We disagree. Furthermore, we believe that the examples presented in this paper substantiate Burgstahler and Jiambalvo's claim that isolating and correcting errors may lead to excessive probability of incorrect acceptance.

4. Some terminology and model assumptions

We summarize some terminology and concepts that can be found in standard textbooks on audit sampling techniques, for example, [**1**]. We shall denote the total book value of the account by M, its unknown error fraction by p, and refer to Mp as the overall error (before correction). In formulas the materiality is denoted by mat; in the examples we used $mat = 0.04M$, that is, 4%. Samples are of size n, and items are sampled proportional to size in Euros (a so-called *monetary unit sample*, a standard auditing practice). For the analysis the all-or-nothing principle is used: this is simplest to model, yet considered sufficient for the investigation at hand (see also Section 9). As a consequence the situation can be considered as if a sample of n Euros is drawn from a population of M Euros that contain a fraction p of 'bad' Euros. Let $J = M/n$; this is the *interval*: the amount represented by one sampled Euro.
Confidence bounds used are 95% upper confidence bounds based on the Poisson approximation to the binomial distribution:

Number of errors in the sample:	0	1	2
Confidence bound on total error:	$3J$	$4.75J$	$6.3J$

Taking $mat = 0.04M$ and the confidence level of 95% are fairly arbitrary choices that are immaterial for the patterns that emerge from the results.

5. Our approach

We consider the auditing process as a testing procedure with null hypothesis "overall error *does* exceed the materiality." This test may

be performed at significance level 5% by computing a 95% upper confidence bound for the overall error and comparing it with the materiality, rejecting the null hypothesis if the bound falls below. In this way a type I error—approving an account with a material error in it—occurs when a sample from such an account yields a confidence bound below the materiality.

An auditor may (and in practice often does) choose the sample size n by solving $3J = mat$, leading to $n = 3M/mat$. If the sample turns out to be free of errors, the upper bound equals the materiality and the account may be approved. If the auditor wants to be able to tolerate one error, then $n = 4.75M/mat$ is chosen.

When isolation and correction is added to this procedure, an account with a material error in it gets a *second chance to slip through*. Its first chance: produce no errors in the sample; this happens with probability at most 5 percent (we note that for an account whose error just exceeds the materiality and a minimal sample size as described above, this probability equals 5 percent). The second chance: produce a pattern of errors whose isolation and correction produces an upper bound below the materiality but is insufficient to bring the overall error below the materiality.

While this reasoning shows that isolation and correction are wrong in the sense that the nominal type I error probability is exceeded, the practical question is by how much this probability is exceeded and whether it can get 'really bad.'

We took as our assignment to find examples where the actual type I error probability *substantially exceeds* its nominal value. We have constructed situations where we can

(1) model what an accountant might do, and
(2) show $\mathrm{P}(\textit{(corrected) account is approved with material error in it}) \geq 0.05$.

A modelling difficulty arose in connection with the 'strata': the segmentation of the account into strata depends on what kind of errors are discovered in the sample. We found it very difficult to describe this in a suitably general model that would permit a probabilistic analysis. We choose to pretend that there exists some stratification of the account (that is, before the sampling is done) and that examination of sampled 'bad' items uncovers (some of) these strata (in Section 6 the strata are known beforehand).

If the account is divided into k strata, we denote their respective book amounts, error fractions and sample sizes by M_i, p_i and n_i, $i = 1, \ldots, k$. They satisfy the relations $M_1 + \cdots + M_k = M$, $M_1 p_1 + \cdots + M_k p_k = Mp$ and $n_1 + \cdots + n_k = n$. In examples most of these parameters will vary, but the sample will always be homogeneously divided among the

strata, i.e., the sampling intervals are all equal: $M_i/n_i = J_i = J$, for all $i = 1, \ldots, k$.
In the next three sections examples are presented of accounts combined with auditing steps that involve isolation and correction. The main difference is in the liberties available and/or taken in this process. In each case we compute the probability of approving an account whose total error (when applicable: after correction) exceeds the materiality.

6. Pre-stratification, errors cannot be corrected

Consider an account with two strata that are delineated *before* the sampling takes place—one may think of 'personnel' and 'materials'. The book sizes are M_1 and M_2, the error fractions p_1 and p_2. So $M = M_1 + M_2$ and the overall error is $Mp = M_1p_1 + M_2p_2$. Errors can only be identified, but not corrected.
The following auditing procedure is followed. A sample of $n = n_1 + n_2$ items is taken, with $n \geq 3M/mat$. If no errors are found, the account is approved. If one or more errors are found, the strata are considered separately. For an error-free stratum the Poisson upper bound is computed: $3M_i/n_i$. A stratum with errors is examined exhaustively and the value of the true error M_ip_i is determined. The (bounds on the) errors are combined by adding them. If this combined bound is below the materiality, then the account is approved.
Our main interest is in the probability of approval when the true error exceeds the materiality, hence we assume $M_1p_1 + M_2p_2 \geq mat$. Let $z_i = (1 - p_i)^{n_i}$, the probability that no errors are found in stratum i. We find:

$$\begin{aligned} \mathrm{P}(\mathit{approval}) = z_1z_2 &+ z_1(1 - z_2)1_{[3M_1/n_1 + M_2p_2 \leq mat]} \\ &+ z_2(1 - z_1)1_{[3M_2/n_2 + M_1p_1 \leq mat]}, \end{aligned}$$

where $1_{[A]} = 1$ when condition A is fulfilled and $1_{[A]} = 0$ otherwise. Intuition suggests that this approval probability should be a decreasing function of both p_1 and p_2. We rewrite the previous expression as the sum of two terms:

$$\begin{aligned} \mathrm{P}(\mathit{approval}) = z_1(1 - z_2)&1_{[3M_1/n_1 + M_2p_2 < mat]} \\ &+ z_2\left(1_{[3M_2/n_2 + M_1p_1 < mat]} + z_1 1_{[3M_2/n_2 + M_1p_1 \geq mat]}\right). \end{aligned}$$

The first term is decreasing in p_1, since z_1 is a decreasing function of p_1. The sum enclosed by square brackets may equal 1 for small values of p_1, and for larger values it equals z_1, hence is a decreasing function of p_1 as well. Thus, we have shown that $\mathrm{P}(\mathit{approval})$ is a decreasing function of p_1 and, by symmetry, also of p_2. Hence, the largest possible approval probabilities occur when the true error $M_1p_1 + M_2p_2$ equals the materiality.

When $n = 3M/mat$ the assumption $\frac{M_i}{n_i} = \frac{M}{n}$ implies that $M_1p_1 + M_2p_2 = mat$ is equivalent with $n_1p_1 + n_2p_2 = 3$. Along this border segment of cases $\mathrm{P}(\mathit{approval}) = z_1z_2$. (Strictly speaking, if p_1 or p_2 is zero we get an extra term; for example, if $p_2 = 0$ the result is $\mathrm{P}(\mathit{approval}) = z_1$; but this corresponds with the limit of $\mathrm{P}(\mathit{approval}) = z_1z_2$ for p_2 going to zero, hence z_2 to 1). This shows that to find the highest approval probability, we should maximize

$$z_1z_2 = (1-p_1)^{n_1}(1-p_2)^{n_2}$$

under the conditions

$$n_1p_1 + n_2p_2 = 3, \quad p_1 \geq 0, p_2 \geq 0.$$

The maximum cannot exceed 0.05 since

$$(1-p_1)^{n_1}(1-p_2)^{n_2} \leq (1-p)^n \leq \mathrm{e}^{-np} = \mathrm{e}^{-3} \leq 0.05,$$

where the first inequality follows by taking logarithms and then applying Jensen's inequality (note that $np = n_1p_1 + n_2p_2$, since $p = mat/M$), and the second from $1 - x \leq \mathrm{e}^{-x}$. The bound 0.05 is attained for $p_1 = p_2 = p$; when p_1 and p_2 are very different the probability is much smaller than 0.05.
When $3M/mat < n \leq 6M/mat$, a similar analysis shows that the approval probability stays below 0.05. For $n > 6M/mat$ it can be exceeded.

In the case of $k \geq 2$ strata, assuming that $Mp \geq mat$, we see that approval occurs when

$$3J + \sum_{i:\text{error found in stratum } i} M_ip_i < mat,$$

where the first term is the error bound for the aggregation of strata where no errors were found, and the second the actual total error for the others. We now restrict ourselves to the special case of 'homogeneous' strata which, admittedly, appears to be the worst case: $p_i = p$, $M_i = M/k$, $i = 1, \ldots, k$. Substituting this we find that approval occurs when the number of strata with errors does not exceed:

$$\frac{mat - 3J}{Mp/k}.$$

Hence, if $mat = 3J$, approval only occurs when no errors are found, and the type I error probability is at most 0.05. This is still the case when the gap $mat - 3J$ is positive but below Mp/k, which is the total error *per stratum*.
However, if the sample size is chosen larger than the minimum $3M/mat$, approval occurs when errors are found in one or more strata. Table 1 lists the probability as a function of the number of strata k. The parameters are $p = 0.05$ and n is chosen to satisfy $mat = 4.75M/n$ as

closely as possible, taking into account that the number sampled per stratum, $n_s = n/k$, should be an integer. Figure 1 shows a plot of these results. Note that in all cases a material error remains, since errors found cannot be corrected. It is seen that the probability of incorrect approval quickly and considerably exceeds acceptable levels as the number of strata increases.

TABLE 1. Probability of incorrect approval for $2, \ldots, 19$ strata; $p = 0.05$.

k	n_s	n	P(*approval*)	k	n_s	n	P(*approval*)
2	60	120	0.002	11	11	121	0.228
3	40	120	0.002	12	10	120	0.223
4	30	120	0.033	13	10	130	0.166
5	24	120	0.028	14	9	126	0.363
6	20	120	0.025	15	8	120	0.394
7	17	119	0.115	16	8	128	0.329
8	15	120	0.102	17	7	119	0.591
9	14	126	0.078	18	7	126	0.528
10	12	120	0.089	19	7	133	0.468

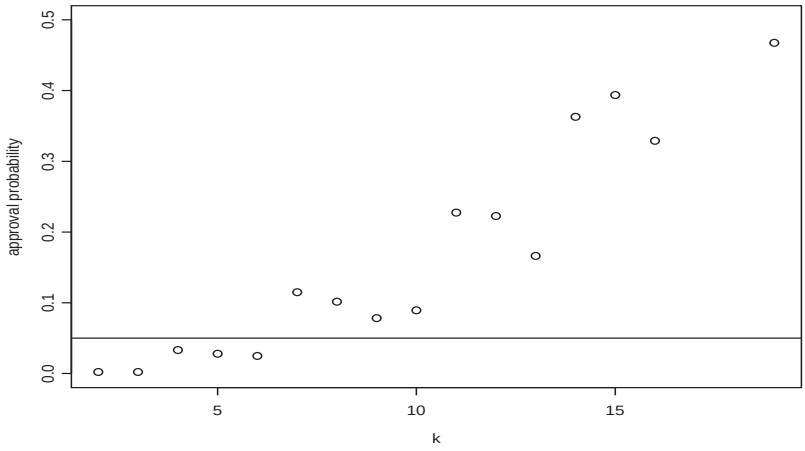


FIGURE 1. Probability of incorrect approval as a function of the number of strata k; $p = 0.05$; horizontal line marks 0.05.

7. Pre-stratification, no correction, dependence on p

In the previous section the situation was analyzed where the number of strata in the pre-stratification varied. In this section we consider the

dependence on the error fraction p. Suppose an auditor judges every error found to be so unique that no others of its kind can be imagined. In effect, this implies that every item is in a stratum on its own and, consequently, there is a large number of strata. Suppose r errors are found in the sample, and error i corresponds to an item of size s_i which after exhaustive examination reveals b_i 'bad' Euros.
Then, considering the part of the sample not contained in the above examined strata as an errorfree sample of $n - r$ from the rest of the account the following (95%) upper confidence bound may be proposed:

$$3\,\frac{M - \sum s_i}{n - r} + \sum b_i$$

For example, let us consider an account with $M = €\,1\,000\,000$ and $k = 1000$ strata, each of size €1000. We use $n = 4M/mat = 100$ and each stratum has the same error fraction p. Then the number of fully examined strata S is approximately binomially distributed with parameters n and p; some strata may be sampled more than once, so the number of fully examined strata is stochastically smaller than this binomial approximation. The actual acceptance probability therefore is bounded below by

$$\mathrm{P}\left(3\frac{M - SM/k}{n - S} + Sp\frac{M}{k} < mat\right).$$

Figure 2 shows a plot of this probability, which is close to 1 for $p \leq 0.11$ and drops below 0.05 only when $p > 0.244$. We see that this case, which could be labeled 'extreme isolation,' produces extremely high probabilities of incorrect approval.

8. Homogenous strata, every error can be corrected

Until now we have considered examples without possibility of correction of errors. In this section we consider the situation where *every* error can be corrected. Suppose an account consists of k homogenous strata of equal size: $p_i = p$, $M_i = M/k$, $i = 1, \ldots, k$. We choose $p = \frac{k}{k-1} \cdot \frac{mat}{M}$, so that correcting one whole stratum still leaves enough errors in the $k - 1$ remaining to attain materiality. This implies the following:

- if zero errors are found in the sample, the account is approved;
- if all errors lie in one stratum, that stratum will be corrected and the resulting account is approved;
- if more than one stratum is corrected the remaining error drops below the materiality.

Let S be the number of strata in which errors are detected. Then S has a binomial distribution with parameters k and $\pi = 1 - (1 - p)^{n_s}$, where

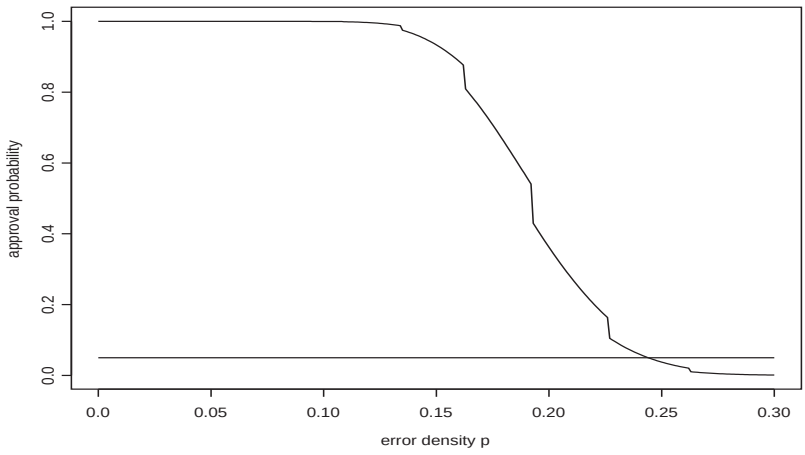


FIGURE 2. Approval probability as a function of the error fraction p; horizontal line marks 0.05.

n_s is the number of items sampled per stratum, equal to n/k rounded upward. It follows that

$$\mathrm{P}(\textit{incorrect approval}) = \mathrm{P}(S \le 1)\,.$$

We have chosen $n \approx 100$, that is, $n \approx 4M/mat$, assuming a sample size slightly above the minimum. In Table 2 and Figure 3 the results are presented. While the results may look less than dramatic, we remark that for the minimum sample size $n = 75$ approval probabilities are larger: for $k = 3, \ldots, 19$, it varies from 0.11 to 0.18. Also, it is suspected that similar examples could be constructed with even larger approval probabilities.

TABLE 2. Probability of incorrect approval for $2, \ldots, 19$ strata; $p = 0.04 \cdot k/(k-1)$.

k	n_s	n	p	P(*approval*)	k	n_s	n	p	P(*approval*)
2	50	100	0.080	0.031	11	10	110	0.044	0.051
3	34	102	0.060	0.041	12	9	108	0.044	0.056
4	25	100	0.053	0.053	13	8	104	0.043	0.065
5	20	100	0.050	0.059	14	8	112	0.043	0.050
6	17	102	0.048	0.059	15	7	105	0.043	0.064
7	15	105	0.047	0.055	16	7	112	0.043	0.051
8	13	104	0.046	0.059	17	6	102	0.042	0.072
9	12	108	0.045	0.053	18	6	108	0.042	0.059
10	10	100	0.044	0.072	19	6	114	0.042	0.048

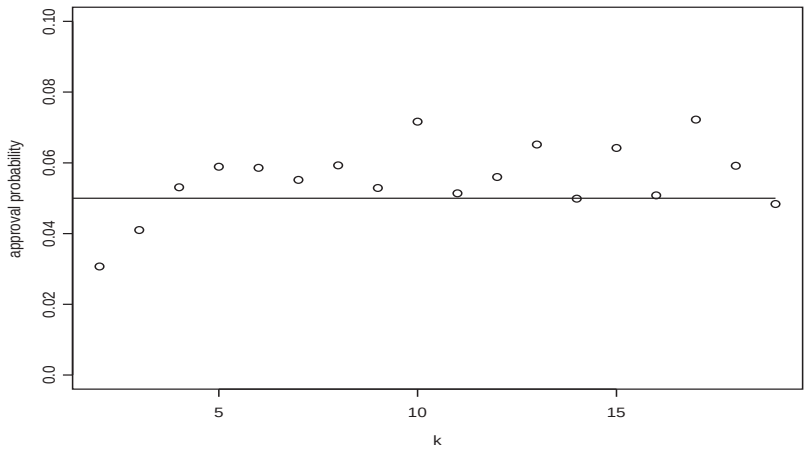


FIGURE 3. Probability of incorrect approval as a function of the number of strata k; $p = 0.04 \cdot k/(k-1)$; horizontal line marks 0.05.

9. Conclusions and final remarks

We briefly discussed two papers dealing with isolating and correcting errors. Burgstahler and Jiambalvo [**2**] predict that isolation and correction may lead to very high probabilities of incorrect approval, whereas Wheeler *et al.* ([**6**]) hold an opposite point of view.
It seems to be difficult to model the isolation and correction procedure used in the auditing process. In order to formulate and analyse a model for this procedure, we have considered three examples with prestratification. For these examples we analysed the effect on the probability of incorrect approval of the number of strata, the error fraction and the possible corrections.
These examples seem to show that in many cases the probability of incorrect approval is (much) larger than the allowed margin, supporting the view of Burgstahler and Jiambalvo.
Using so-called Stringer bounds will not amend this situation: we believe that similar examples with similar excessively high approval probabilities can be constructed. We conclude that reconsideration of the standards pertaining to isolation and correction, as formulated in the *International Standard on Auditing 530*, seems to be in order.

Acknowledgements

We would like to thank Ed Broeze and Berrie Zielman of the Netherlands Court of Audit (de Algemene Rekenkamer) for introducing us to this problem and for helpful and stimulating discussions.

References

[1] A.A. Arens and J.K. Loebbecke. *Auditing: an integrated approach*. Prentice-Hall, Upper Saddle River, 1998.

[2] D. Burgstahler and J. Jiambalvo. Sample error characteristics and projection of error to audit populations. *The Accounting Review*, 61(2):223–248, 1986.

[3] R.B. Dusenberg, J.L. Reimers, and S.W. Wheeler. The effect of containment information and error frequency on projection of sample errors to audit populations. *The Accounting Review*, 69(1):257–264, 1994.

[4] R.J. Elder and R.D. Allen. An empirical investigation of the auditor's decision to project errors. *Auditing: A Journal of Practice & Theory*, 17(2):71–87, 1998.

[5] International Federation of Accountants. *Handbook on International Auditing, Assurance and Ethics Pronouncements*. IFAC, 2004.

[6] S.W. Wheeler, R.B. Dusenberg, and J.L. Reimers. Projecting sample misstatements to audit populations: Theoretical, professional, and emperical considerations. *Decision Sciences*, 28(2):261–278, 1997.

Statistical Disclosure Control using PRAM

Eric Cator, André Hensbergen†, Yves Rozenholc‡

† TU Delft, EWI (DIAM), Mekelweg 4, 2628 CD Delft, e.a.cator@ewi.tudelft.nl, a.t.hensbergen@ewi.tudelft.nl
‡ Université Pierre et Marie Curie - Boîte courrier 188, P.O.Box 75252 Cedex 05 Paris France, yves.rozenholc@math.jussieu.fr

Abstract. We will look at Statistical Disclosure Control where our goal is to protect microarray data against spontaneous recognition by a user, by applying a known transition matrix P to each row of our microarray. This method was invented by Statistics Netherlands and is called PRAM. We want to choose the transition matrix P in such a way that our loss of information is minimal, while at the same time guaranteeing a certain level of security. In this paper we will define what is meant by this level of security and we will consider possible loss of information measures, also focussing on the applicability of the method.

1. Introduction

Statistics Netherlands (CBS) gathers each year a huge amount of data concerning the Dutch people. Some of these data are sensitive, and CBS is only able to obtain these data if it can guarantee that this sensitive data is not directly available to outsiders. Also Dutch privacy legislation is quite strict and enforces CBS to take appropriate measures. We will focus on one of these measures, namely to ensure that if CBS lets some researcher or company (we will call this the user) work with certain microarray data, then CBS has to modify this microarray in such a way that each individual in the data (this corresponds to one row in the microarray) is protected against spontaneous recognition by the user. This definition is still rather vague, so let us describe the situation in more detail.

Our microarray consists of a matrix, where each row corresponds to some individual (this maybe for example a person or a company) and each column corresponds to some property of the individual, for example age, salary of place of residence, which we will call the variables. CBS distinguishes two kinds of variables: the identifying variables and the sensitive variables. The identifying variables are properties of the individual that others may know or could easily find out, such as sex or place of residence. The sensitive variables are properties of the individual that are normally not known to outsiders, such as salary or

health. We will call the identifying variables $\xi_1, \xi_2, \ldots, \xi_m$ and the sensitive variables $y_1, \ldots, y_M$. We define $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_m$ the space of possible values for $\xi = (\xi_1, \ldots, \xi_m)$. Likewise we define $\mathcal{Y}$ as the space of possible values for $y = (y_1, \ldots, y_M)$. Furthermore, we will call the number of individuals in the data n. We use the notation $\xi_j(i)$ for the value of the j^{th} variable corresponding to individual i. Similarly we define $\xi(i)$ and $y(i)$.
CBS states that when a user spontaneously recognizes an individual, she can only do this by overviewing 3 identifying variables at the same time. If she were to use more than 3 variables, she would be intentionally trying to find someone, and CBS has taken legal measures to prevent this. This choice of three identifying variables is defined by choosing a map

$$\pi : \{1, 2, 3\} \to \{1, 2, \ldots, m\}.$$

The set of all possible choices of three variables that we want to protect against is denoted by Π. We define $\xi_\pi = (\xi_{\pi(1)}, \xi_{\pi(2)}, \xi_{\pi(3)})$, and likewise $\mathcal{X}_\pi$. In Section 2 we will define precisely what we mean by protecting against spontaneous recognition.
The way we will secure our data is by using a transition matrix P_{kl}, where $k, l \in \mathcal{X}$. This means that the actual data set CBS will give to its user consists of entries $(X(i); y(i))$ $(i = 1, \ldots, n)$, where each $X(i)$ is a random variable such that

$$\mathrm{P}\left(X(i) = l \mid \xi(i) = k\right) = P_{kl}.$$

Each row will be transformed like this, independently of each other, so all $X(i)$ are independent, but not identically distributed! This method of securing the data is called PRAM. The idea is that the user not only receives the modified data, but also the transition matrix P, so that she can still make proper statistical inference, but she cannot spontaneously recognize an individual and thus find out sensitive information about him. In Section 3 we will describe how we can choose P such that the data is secure, but the loss of information is minimal.

2. Secure against spontaneous recognition

Suppose our user received the microarray $(x(i); y(i))$ $(i = 1, \ldots, n)$ from CBS, together with a transition matrix P. She would normally try in some way to retrieve the original data (ξ, y), then look at three of the identifying variables ξ_π and think that she recognizes a certain individual. She knows that this individual, let's call him John, has the value $k_0 \in \mathcal{X}_\pi$ for the three identifying variables $\pi(1), \pi(2)$ and $\pi(3)$. She is therefore interested in the possibility that $\xi_\pi(i) = k_0$, for a certain i that she believes to be John. However, this is not all. If she concludes

that indeed $\xi_\pi(i) = k_0$, then she still might not be sure that i is John. It might be that a lot of individuals have the value k_0 for ξ_π.
Let us suppose our user takes the following Bayesian approach in deciding whether i is John or not: since there are n individuals in the microarray, the prior probability (so without any information) of i being John equals $1/n$. Now she knows that $x_\pi(i) = k_0$. Therefore we get

$$\begin{aligned}
\mathrm{P}\left(i = \text{John} \mid x_\pi(i) = k_0\right) &= \frac{\mathrm{P}\left(x_\pi(i) = k_0 \mid i = \text{John}\right) \cdot \mathrm{P}(i = \text{John})}{\sum_{j=1}^{n} \mathrm{P}\left(x_\pi(j) = k_0 \mid \xi(j)\right) \cdot \frac{1}{n}} \\
&= \frac{\sum_{k:k_\pi = k_0} P_{\xi(\text{John})\mathrm{k}}}{\sum_{j=1}^{n} \sum_{k:k_\pi = k_0} P_{\xi(j)k}} \\
&= \frac{\sum_{k:k_\pi = k_0} P_{\xi(\text{John})\mathrm{k}}}{\sum_{l \in \mathcal{X}} U_0(l) \sum_{k:k_\pi = k_0} P_{lk}}.
\end{aligned}$$

Here we define

$$U_0(l) = \#\{j : \xi(j) = l\}.$$

She would decide that indeed $i =$John if

$$\mathrm{P}\left(i = \text{John} \mid x_\pi(i) = k_0\right) > \alpha,$$

for some significance level α (which may be specified by CBS). Therefore, if we want to protect our data, we need the following condition on the transition matrix P:

C1: For each $m \in \mathcal{X}$ such that there exists i with $\xi(i) = m$, for each $\pi \in \Pi$ and $k_0 \in \mathcal{X}_\pi$, we must have

$$\frac{\sum_{k:k_\pi = k_0} P_{mk}}{\sum_{l \in \mathcal{X}} U_0(l) \sum_{k:k_\pi = k_0} P_{lk}} \leq \alpha. \tag{1}$$

For example, if P is the identity matrix (so the original data is given to the user), then for $m \in \mathcal{X}$ with $m_\pi = k_0$, we get

$$\frac{\sum_{k:k_\pi = k_0} P_{mk}}{\sum_{l \in \mathcal{X}} U_0(l) \sum_{k:k_\pi = k_0} P_{lk}} = \frac{1}{U_\pi(k_0)}.$$

Here,

$$U_\pi(k_0) = \#\{j : \xi_\pi(j) = k_0\}.$$

So condition (C1) can only hold if for all π and k_0 we have $U_\pi(k_0) \geq 1/\alpha$ (or $U_\pi(k_0) = 0$). In other words, there should be no rare combinations in the original data.

The other extreme is when P has constant entries, so all information about the data is lost. Then

$$\frac{\sum_{k:k_\pi=k_0} P_{mk}}{\sum_{l\in\mathcal{X}} U_0(l) \sum_{k:k_\pi=k_0} P_{lk}} = \frac{1}{n},$$

so condition (C1) is always satisfied (at least if $n \geq 1/\alpha$). This shows that there always exist transition matrices that satisfy (C1).

One final remark is that

$$\frac{\sum_{k:k_\pi=k_0} P_{mk}}{\sum_{l\in\mathcal{X}} U_0(l) \sum_{k:k_\pi=k_0} P_{lk}} \leq \frac{1}{U_0(m)}. \tag{2}$$

This shows that if $U_0(m) \geq 1/\alpha$, so if $m \in \mathcal{X}$ occurs frequently enough in the original data, then Condition (C1) always holds for this m, for any choice of π, k_0 and P.

Condition (C1) can be rewritten as a linear constraint for P. However, (1) has to hold for all $\pi \in \Pi$ and all $k_0 \in \mathcal{X}_\pi$. The total number of conditions might grow exponentially (or even faster) if one doesn't control the set Π properly. One would have to think about which combinations of variables is reasonable for a spontaneous recognition.

3. Minimal loss of information

We now know which condition the transition matrix P has to satisfy, but we still have to choose an optimal P, in some sense. We want to choose P such that the information loss is minimal, but we need to quantify this loss.
Our user is interested in estimating some property of the original data $\{(\xi(i); y(i)) : 1 \leq i \leq n\}$. In fact, she will be interested in a property of the empirical measure μ of all the rows $(\xi(i); y(i))$. A logical way to estimate this property is using the data available to her (i.e. $\{(x(i); y(i)) : 1 \leq i \leq n\}$) to obtain an estimate $\hat{\mu}$ of μ, and use this $\hat{\mu}$ to estimate this property. Since we do not know which property the user will be interested in, we could try and make sure that $\hat{\mu}$ is a good estimate for μ. We have to choose an estimator for μ, and we choose the Maximum Likelihood estimator. Our information loss will now be measured in terms of how far in expectation $\hat{\mu}$ lies from μ.

The Maximum Likelihood Estimator. To define the MLE, the user will assume that our original data is generated by some measure μ^* on $\mathcal{X} \times \mathcal{Y}$. So she assumes that

$$(\Xi; Y) \sim \mu^*.$$

Given the transition matrix P, she then knows that the log-likelihood of her data $\{(x(i); y(i)) : 1 \leq i \leq n\}$ is given by:

$$l(\mu^*) = \sum_{i=1}^{n} \log(\sum_{k \in \mathcal{X}} \mu^*(k; y(i)) \cdot P_{kx(i)}).$$

Now define $\mu^*_{\mathcal{Y}}$ as the marginal measure on $\mathcal{Y}$, i.e. the distribution of Y. Then

$$\mu^*(k; y) = \mathrm{P}^*(\Xi = k | Y = y) \cdot \mu^*_{\mathcal{Y}}(y).$$

Define

$$\mu^*(k|y) = \mathrm{P}^*(\Xi = k | Y = y).$$

We then have

$$l(\mu^*) = \sum_{i=1}^{n} \log(\mu^*_{\mathcal{Y}}(y(i))) + \sum_{i=1}^{n} \log(\sum_{k \in \mathcal{X}} \mu^*(k|y(i)) \cdot P_{kx(i)}).$$

This shows that we can maximize over $\mu^*_{\mathcal{Y}}$ and $\mu^*(k|y)$ separately. The first term is just the log-likelihood when we have a sample $y(1), \ldots, y(n)$ from $\mu^*_{\mathcal{Y}}$, so the MLE for $\mu^*_{\mathcal{Y}}$ is equal to the empirical measure of the $y(i)$'s, which makes sense, since the $y(i)$'s are not changed by P. So $\hat{\mu}_{\mathcal{Y}} = \mu_{\mathcal{Y}}$.
In order to find the MLE for $\mu^*(k|y)$, we introduce some vector notation: P is a matrix in $\mathbb{R}^{\mathcal{X} \times \mathcal{X}}$ and each μ^*_i is a vector in $\mathbb{R}^{\mathcal{X}}$ such that

$$\mu^*_i(k) = \mu^*(k|y(i)).$$

Define the following vectors in $\mathbb{R}^{\mathcal{X}}$:

$$b_i = P^t \mu^*_i.$$

For $a, b \in \mathbb{R}^{\mathcal{X}}$ we define $\langle a, b \rangle = \sum_{k \in \mathcal{X}} a(k)b(k)$. Then

$$\langle b_i, 1 \rangle = \langle \mu^*_i, P1 \rangle = \langle \mu^*_i, 1 \rangle = 1.$$

This means that each b_i represents a probability measure on $\mathcal{X}$. Furthermore, if P is invertible, which we will assume from now on, we get that

$$\mu^*_i = (P^{-1})^t b_i.$$

So if we can maximize

$$\tilde{l}(b) = \sum_{i=1}^{n} \log(b_i(x(i))), \tag{3}$$

we would have that

$$\hat{\mu}(k|y(i)) = \sum_{l \in \mathcal{X}} \hat{b}_i(l) P^{-1}_{lk}.$$

However, in Equation (3) we recognize the loglikelihood for the data $\{x(i) : 1 \leq i \leq n\}$ with

$$\mathrm{P}(X(i) = k) = b_i(k).$$

So if we define

$$U(k; y) = \#\{i : (x(i); y(i)) = (k; y)\} \quad \text{and} \quad U(\cdot; y) = \#\{i : y(i) = y\},$$

we get that

$$\hat{b}_i(k) = \frac{U(k; y(i))}{U(\cdot; y(i))}.$$

Our conclusion is (note that $\mu_{\mathcal{Y}}(y) = U(\cdot; y)/n$):

$$\hat{\mu}(k; y) = \frac{1}{n} \sum_{l \in \mathcal{X}} P^{-1}_{lk} U(l; y). \tag{4}$$

Note that we can interpret $U(k; y)$ as a random variable by replacing $x(i)$ by $X(i)$ in its definition; this also makes $\hat{\mu}$ into a random measure, so we can calculate its expected distance (which we need to choose) to μ.

The L^2 loss. We define the loss as some expected deviation of $\hat{\mu}$ from μ, the original empirical measure. We will take this expectation only with respect to the change of the variables using P, so we will no longer view $(\xi(i); y(i))$ as a realization of a random variable. In information theory, a usual measure for the deviation of $\hat{\mu}$ from μ is the Kullback-Leibler divergence:

$$l_{\mathrm{KL}}(P) = \mathrm{E}_P \left[\sum_{(k;y) \in \mathcal{X} \times \mathcal{Y}} - \log \left(\frac{\hat{\mu}(k; y)}{\mu(k; y)} \right) \mu(k; y) \right].$$

Here $\mathrm{E}_P[\cdot]$ denotes expectation with respect to the probability measure induced by the transition matrix P; remember that $\hat{\mu}(k; y)$ is a stochastic variable whose distribution depends on P (and on the original data, of course). This is why we denote our loss function l_{KL} as a function of P.
Another possible choice for our loss function is the quadratic loss, or L^2-loss:

$$l_2(P) = \mathrm{E}_P \left[\sum_{(k;y) \in \mathcal{X} \times \mathcal{Y}} \left(\hat{\mu}(k; y) - \mu(k; y) \right)^2 \right].$$

This leads to a more feasible loss function than l_{KL}, i.e. one that leads to an easier optimization problem, which is important if we want the method to work for large datasets and, more importantly, for large Π.

We will spend some time evaluating $l_2(P)$. Define

$$U_0(l;y) = \#\{i : (\xi(i);y(i)) = (l;y)\}.$$

We will assume that P is invertible. This means that we can use Equation (4) to see that

$$\begin{aligned}
\mathrm{E}_P[\hat{\mu}(k;y)] &= \frac{1}{n}\sum_{l\in\mathcal{X}} P^{-1}_{lk}\mathrm{E}_P[U(l;y)] \\
&= \frac{1}{n}\sum_{l\in\mathcal{X}} P^{-1}_{lk}\sum_{l'\in\mathcal{X}} P_{l'l}U_0(l';y) \\
&= \frac{1}{n}U_0(k;y) \\
&= \mu(k;y).
\end{aligned}$$

This shows that $\hat{\mu}$ is an unbiased estimator of μ. So we get

$$l_2(P) = \sum_{(k;y)} \mathrm{E}_P\left[\hat{\mu}(k;y)^2\right] - \mu(k;y)^2.$$

Now we concentrate on the random variables $U(l;y)$, for fixed y:

$$\begin{aligned}
\mathrm{E}_P[U(l;y)U(l';y)] &= \mathrm{E}_P\left[\sum_{\{i,j:y(i)=y(j)=y\}} 1_{\{X(i)=l\}}1_{\{X(j)=l'\}}\right] \\
&= \sum_{\{i\neq j:y(i)=y(j)=y\}} P_{\xi(i)l}P_{\xi(j)l'} + \delta_{ll'}\sum_{\{i:y(i)=y\}} P_{\xi(i)l}.
\end{aligned}$$

This means that

$$\begin{aligned}
n^2\mathrm{E}_P\left[\hat{\mu}(k;y)^2\right] &= \sum_{l,l'\in\mathcal{X}} P^{-1}_{lk}P^{-1}_{l'k}\mathrm{E}_P[U(l;y)U(l';y)] \\
&= \sum_{l,l'\in\mathcal{X}}\sum_{\{i\neq j:y(i)=y(j)=y\}} P^{-1}_{lk}P^{-1}_{l'k}P_{\xi(i)l}P_{\xi(j)l'} + \sum_{l\in\mathcal{X}}\sum_{\{i:y(i)=y\}}\left(P^{-1}_{lk}\right)^2 P_{\xi(i)l} \\
&= \sum_{\{i\neq j:y(i)=y(j)=y\}} \delta_{\xi(i)k}\delta_{\xi(j)k} + \sum_{l\in\mathcal{X}}\sum_{\{i:y(i)=y\}}\left(P^{-1}_{lk}\right)^2 P_{\xi(i)l} \\
&= U_0(k;y)^2 - U_0(k;y) + \sum_{l\in\mathcal{X}}\sum_{\{i:y(i)=y\}}\left(P^{-1}_{lk}\right)^2 P_{\xi(i)l}.
\end{aligned}$$

Therefore,

$$\begin{aligned} l_2(P) &= \sum_{(k;y)\in\mathcal{X}\times\mathcal{Y}} \left[\frac{U_0(k;y)^2}{n^2} - \frac{U_0(k;y)}{n^2} + \frac{1}{n^2} \sum_{l\in\mathcal{X}} \sum_{\{i:y(i)=y\}} \left(P^{-1}_{lk}\right)^2 P_{\xi(i)l} - \frac{U_0(k;y)^2}{n^2} \right] \\ &= \frac{1}{n^2} \sum_{(k;y)\in\mathcal{X}\times\mathcal{Y}} \sum_{\{i:y(i)=y\}} \sum_{l\in\mathcal{X}} \left(P^{-1}_{lk}\right)^2 P_{\xi(i)l} - \frac{1}{n} \\ &= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{k,l\in\mathcal{X}} \left(P^{-1}_{lk}\right)^2 P_{\xi(i)l} - \frac{1}{n}. \end{aligned}$$

This function of P is feasible as a loss function, as we will show in the next section.

4. Implementation

We will focus on a relatively small example, where $\mathcal{X} = \{0,1\}^4 \times \{0,1,2\}$, with 48 elements. This means that our PRAM matrix P will be a 48×48 matrix, where 48 is the total number of distinct elements in $\mathcal{X}$. It is easy to check that the number of different combinations (π, k_0) is 104. Handling bigger data sets requires a more sophisticated approach to the optimization problem, but we do wish to point out that according to inequality (2), we only need to check Condition (C1) for those $m \in \mathcal{X}$ whose frequency in the micro array is smaller than $1/\alpha$, where α is the required security level. This means that if we consider a lot of identifying variables (so $\mathcal{X}$ is big), but our micro array contains a lot of individuals, we still might end up with a reasonably small number of side conditions for our optimization. We will give some comments on bigger $\mathcal{X}$ later on.

In our example we used a micro array of 2500 individuals and a security level of $\alpha = 0.1$. Furthermore we incorporated some probability structure on the variables to get a reasonable number of rare individuals. It turned out that seventeen individuals m from $\mathcal{X}$ satisfied $0 < U_0(m) < 10$ $(= 1/\alpha)$. Only for those seventeen we needed to check Condition (C1). A good starting value for P when solving the optimization problem is the identity matrix I. In general, for large micro arrays it turns out that the optimal P will be very close to I. The explanation is: suppose we give elements in $\mathcal{X}$ that are quite frequent in the micro array small probabilities to be changed into rare elements, and leave the other elements unaltered. Then it will already be likely that the occurrence of a rare combination of π and k_0 is actually due to the PRAM transformation, and therefore will not lead to the identification of an individual. Of course, if there are many rare elements in $\mathcal{X}$ and few frequent ones, the optimal P might be quite different from I.

To check that the optimal P in our example is indeed close to I, we noted that the lowest diagonal element was 0.9971. Furthermore, we found that, applying the PRAM transformation to our array X, using this P, on average leads to only 4 changes (out of 2500 possible changes). This implies only a small loss in information, and indeed the L^2-loss is a mere $6.4 \cdot 10^{-7}$. To give some more insight into the structure of P, we consider Figure 1, which shows the probability of an element being altered, set out against its frequency in the micro array X.

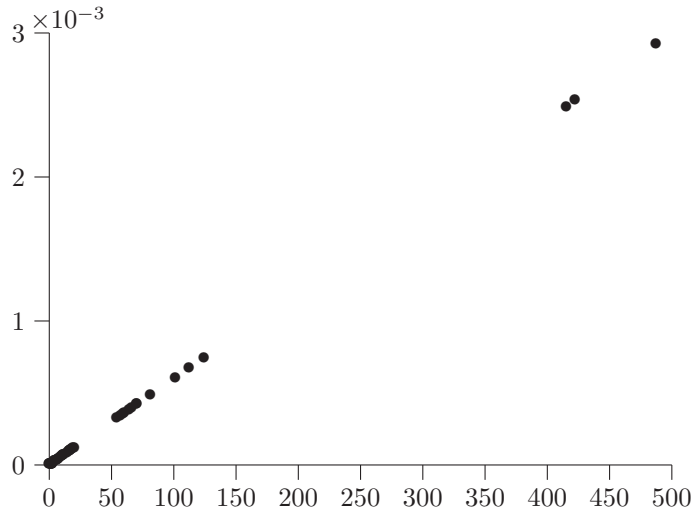


FIGURE 1. Probability of change against frequency.

As one can see, there is almost a linear relationship. When zooming in it appears that only for elements with very small frequencies there are deviations from this linearity.

We also noted that if an element is changed during the PRAM procedure, there is a 0.99999 chance that it gets changed into one of the following three elements: $(0,0,0,0,0)$, $(0,0,0,0,1)$ or $(0,0,0,0,2)$. The reason is that there were only five combinations of (π, k_0) that satisfied $U_\pi(k_0) < 1/\alpha$, and for all these five combinations there exists a k out of the three elements above, such that $k_\pi = k_0$. In other words, these three elements are enough to take care of the problem cases. Remember that only if there are (π, k_0) with $0 < U_\pi(k_0) < 1/\alpha$ something has to be done, that is, only then the micro array needs to be changed (otherwise Condition (C1) would be satisfied for $P = I$, the identity matrix.) Finally, it turned out that when the PRAM transformation is applied to the original micro array, in 88 % of the cases an element that undergoes a change is one of the three most frequently occurring elements.

These observations can help when dealing with large micro arrays: to find a good PRAM matrix, it might be enough to optimize over all P that only allow changes from frequent elements to elements that together cover the problem combinations of π and k_0. This way one may not find the optimal P, but perhaps one whose information loss is still acceptable.

We conclude with two inequalities containing the left hand side of (C1) that are useful for big micro arrays. For both inequalities we assume that the PRAM matrix P has all its diagonal elements $P_{kk} > \gamma$, for some γ close to 1.

First, for any $m \in \mathcal{X}$:

$$\begin{aligned}\frac{\sum_{k:k_\pi=k_0} P_{mk}}{\sum_{l\in\mathcal{X}} U_0(l) \sum_{k:k_\pi=k_0} P_{lk}} &\leq \frac{1}{\sum_{k:k_\pi=k_0} U_0(k) P_{kk}} \\ &\leq \frac{1}{\gamma U_\pi(k_0)}.\end{aligned}$$

This means that we only have to check combinations of π and k_0 that satisfy

$$U_\pi(k_0) \leq \frac{1}{\gamma\alpha}.$$

In our example we could have picked $\gamma = 0.99$, since our optimal P has all diagonal elements bigger than 0.99, but we do not know this before we do the optimization. However, choosing $\gamma = 0.9$ could already drastically reduce the number of combinations that needs to be checked. If in this way you find diagonal elements lower than γ, you could start the optimization again with a smaller γ.

Secondly, for elements $m \in \mathcal{X}$ for which $m_\pi \neq k_0$, we can show that

$$\frac{\sum_{k:k_\pi=k_0} P_{mk}}{\sum_{l\in\mathcal{X}} U_0(l) \sum_{k:k_\pi=k_0} P_{lk}} \leq \frac{1-\gamma}{\gamma \cdot U_\pi(k_0)}. \tag{5}$$

Namely, if $m_\pi \neq k_0$, for the numerator we have

$$\sum_{k:k_\pi=k_0} P_{mk} \leq 1-\gamma,$$

since the sum does not contain the diagonal element, and the denominator is taken care of as above. So if the right hand side of (5) is smaller than α, then for the given m condition (C1) is certainly fulfilled. Now from $U_\pi(k_0) \geq 1$ it follows that

$$\frac{1-\gamma}{\gamma \cdot U_\pi(k_0)} \leq \frac{1-\gamma}{\gamma},$$

so if the last expression is smaller than α, we have can verify Condition (C1) by checking all $m \in \mathcal{X}$, $\pi \in \Pi$ and $k_0 \in \mathcal{X}_\pi$ such that

- $U_0(m) < \frac{1}{\alpha}$

- $U_\pi(k_0) < \frac{1}{\gamma\alpha}$
- $m_\pi = k_0$.

The rotor spinning process for fibre production

P. den Decker, H. Knoester, H. Meerman[†], K. Dekker, W. van Horssen, C. Vuik, P. Wesseling[‡], G. Prokert[§], B. van 't Hof[♯], F. van Beckum[♯]

1. Introduction

At Tejin Twaron in Arnhem new ways of producing fibres are being developed. One of the interesting techniques is the so-called Rotor Spinning Process. In principle, this process looks a lot like the making of sugarflos (or cotton candy) at a fair. Here, however, we deal with a polymer-filled disc with tiny holes. The polymer is pressed, due to the centrifugal forces, through the holes to the outside. This process is already in operation at the company; at Tejin Twaron there is also a pilot machine in which variations in the process and geometry can be tested.
The liquid polymer solidifies and becomes a thin filament on the exterior boundary of the machine. The purpose of the work during the week "Mathematics with Industry" is to verify an existing model on the basis of a momentum equation and mass balance and if possible to improve the model.
A first order approximation of the path of the filament (without modelling air friction) in the space between disc and exterior boundary of the machine exists already. Also a description of the path with water cooling and air friction is available. However, the model can be improved: certain states of the rotor spin process should be approximated in a better way. The ultimate purpose of the modelling in more detail reads:

1 Try to describe the situation (process and geometry) in which continuous filaments can be generated. Breaking of filaments may cause problems in the use of the material if the length of the filament is below a critical length.
2 Try to determine the circumstances (process and geometry) in which the length of a broken filament can be determined beforehand. In this case fibres can, in principle, be produced.
3 Determine the effect of processing conditions (e.g. temperature, rotor speed) in the present operating situation in order to achieve a robust production process.

During the workshop we mainly focused on the existing model used at Tejin Twaron, a boundary-value problem for differential equations, which is described in Section 2. As shooting techniques from the NAG-library did not yield acceptable results at Tejin Twaron, we considered the model from various perspectives. One subgroup tried to solve the model numerically, using MATLAB, by Picard-iteration starting from a model without viscosity. A second subgroup analyzed the model using perturbation theory in the neighborhood of the orifice and the last subgroup derived a time-dependent description including an energy balance. Their results are presented in the subsequent sections. Finally we present our conclusions and ideas for further research.

2. The mathematical model

A disc with radius R_{rot} is rotating anti-clockwise with angular velocity ω. The polymer (density ρ and viscosity η) is extruded from an orifice, and subsequently moves in the direction of the coagulator, which has radius R_{coag}. In a stable stationary process the trajectory of the polymer, the so-called spinning line, will be fixed in a rotating coordinate system. Therefore, we omit the time-derivatives in the model, but consider the movement of the polymer along the spinning line (see Fig. 1). Then, the independent variable is s, the arc-length along the

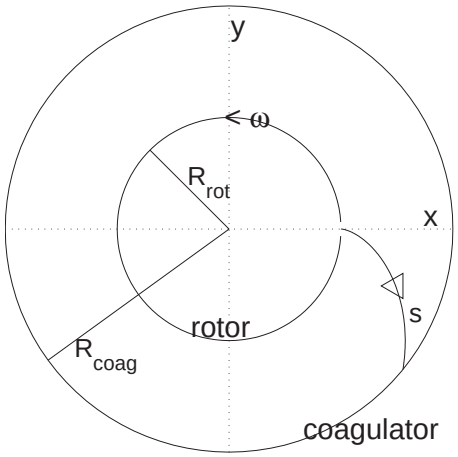


FIGURE 1. Spinning line in rotating coordinate system

spinning line, and the dependent variables are the position x, y of the spinning line, the velocity v of the polymer along the spinning line, and the diameter A of the spinning line. The variables x, y, s satisfy the equation

$$\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 = 1, \tag{1}$$

and the balance of mass gives a relation between A and v, as the mass flux Φ should be constant:

$$\rho A v = \Phi.$$

In the simple model without air friction we consider the viscous forces, the centrifugal force and the Coriolis force acting on an element of the spinning line of size Δs and position x, y:

$$\begin{aligned}
\mathbf{F}_{centr} &= \Delta s\, \rho A \omega^2 \begin{bmatrix} x \\ y \end{bmatrix}, \\
\mathbf{F}_{cor} &= 2\Delta s\, \rho A \omega v \begin{bmatrix} \frac{dy}{ds} \\ -\frac{dx}{ds} \end{bmatrix}, \\
\mathbf{F}_{visc} &= \Delta s \begin{bmatrix} \frac{d}{ds}\left(F\frac{dx}{ds}\right) \\ \frac{d}{ds}\left(F\frac{dy}{ds}\right) \end{bmatrix}.
\end{aligned}$$

Here, F denotes the norm of the viscous force vector at s. Balance of momentum then leads to the second order differential equations

$$(F - \Phi v)\frac{d^2x}{ds^2} = -\frac{\Phi\omega^2 x}{v} - 2\Phi\omega\frac{dy}{ds} - \frac{dx}{ds}\frac{d}{ds}(F - \Phi v), \tag{2}$$

$$(F - \Phi v)\frac{d^2y}{ds^2} = -\frac{\Phi\omega^2 y}{v} + 2\Phi\omega\frac{dx}{ds} - \frac{dy}{ds}\frac{d}{ds}(F - \Phi v). \tag{3}$$

Further it is assumed that the polymer is Newtonian, so the viscous force satisfies

$$\frac{dv}{ds} = \frac{\rho v}{\eta}\frac{F}{\Phi}. \tag{4}$$

Instead of solving the differential equations (2-4), together with the algebraic condition (1), it seemed more appropriate to replace (1) by a differential equation. Taking the inner product of vectorial momentum equations (2-3) and the vector

$$\left(\frac{dx}{ds}, \frac{dy}{ds}\right)^T,$$

using (1) and its differentiated form, leads to

$$\frac{dF}{ds} = \Phi\frac{dv}{ds} - \frac{\Phi\omega^2}{v}\left(x\frac{dx}{ds} + y\frac{dy}{ds}\right). \tag{5}$$

The initial conditions for the system (2-5) are

$$x(0) = R_{rot}, \quad y(0) = 0, \quad v(0) = v_0, \quad F(0) = F_0, \tag{6a}$$

$$\frac{dx}{ds} = 1, \quad \frac{dy}{ds} = 0, \quad \text{for } s = 0. \tag{6b}$$

The viscous force F_0 is unknown and should follow from conditions imposed on the spinning line at the coagulator. As the arc length L of the

spinning line at the coagulator is still unknown, we need two boundary conditions. The first one is given by the radius of the coagulator; moreover the velocity of the spinning line is assumed to be known:

$$x(L)^2 + y(L)^2 = R_{coag}, \quad v(L) = v_e.$$

It is easy to solve these equations in case of zero viscosity. Then, the model reduces to the problem of bullets fired from a rotating disc. Obviously, they move in a straight line, and the solution in a rotating coordinate system is given by

$$x(t) = (R_{rot} + v_0 t)\cos(\omega t) + \omega R_{rot} t \sin(\omega t), \tag{7a}$$

$$y(t) = -(R_{rot} + v_0 t)\sin(\omega t) + \omega R_{rot} t \cos(\omega t), \tag{7b}$$

$$v(t) = \sqrt{(v_0 + \omega^2 R_{rot} t)^2 + (\omega v_0 t)^2}, \tag{7c}$$

where s and t are related by $v = ds/dt$. A solution is presented in Fig. 2.

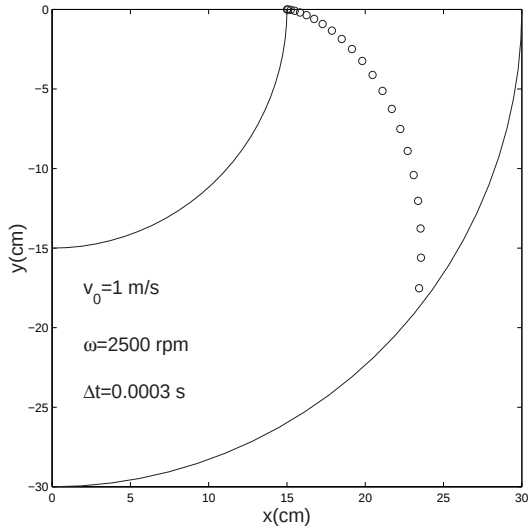


FIGURE 2. Spinning line without viscous forces

Realistic values for the physical parameters are in the viscous case:

$$\rho = 1700 \quad kg/m^3, \qquad \eta = 1200 \quad Pa\,s, \qquad v_0 = 1 \quad m/s,$$

$$R_{rot} = 0.15 \quad m, \qquad R_{coag} = 0.3 \quad m, \qquad A(0) = \pi/64\ 10^{-6} \quad m^2.$$

Originally, it was decided to use shooting techniques for system (2-5) with appropriate choices for F_0 in (6). The differential equations were solved by a routine for stiff systems from the NAG-library. However, difficulties arose as the solution appeared to be extremely sensitive to the choice for the initial viscous force. In many cases the solution became unstable and sometimes the spinning line moved in the wrong direction.

3. Numerical solution methods

The system of 2 second-order and 2 first-order differential equations (2-5) could be discretized as a boundary-value problem on the interval $[0, L]$, where the value of L is yet unknown. This approach, however, would involve some programming effort, resulting in a large system of nonlinear equations, which had to be solved several times to find an appropriate value for L. Moreover, it was not a priori clear how to choose the grid points, as we suspected that an equidistant grid would not do. Therefore we deemed this approach not to be feasible within the limited time available, and beyond that, in case of failure the intrinsic difficulty in the model would not be revealed.

We then decided to split the system into two parts, such that each could be solved in a straightforward manner. The second-order equations for x and y, (2-3) can be solved by an integration method when F and v are known. Once x and y are known, a boundary-value technique could be applied to solve the second-order system for v,

$$\frac{\eta}{\rho}\left(-\frac{d^2v}{ds^2}+\frac{1}{v}\left(\frac{dv}{ds}\right)^2\right)+v\frac{dv}{ds}=\omega^2\left(x\frac{dx}{ds}+y\frac{dy}{ds}\right), \tag{8}$$

which is obtained by substitution of (4) into (5). For simplicity, we impose here a Neumann boundary condition at $s = L$

$$\frac{dv}{ds}=0, \quad (s=L).$$

To start with, we solve the non-viscous model ($\eta = 0$) which has the known solution given by (7). Thereafter, we alternate solving system (2-3) and equation (8) using a Picard-iteration, with the idea that this process might converge to the solution of the complete system (2-5). In doing so, we encountered several problems.

First, observe that the second derivatives in (2-3) are multiplied by the factor $F - \Phi v$, so the system becomes singular whenever this factor changes sign, and the solution will explode. This phenomenon does not occur for small values of the viscosity η, as F is then small too. However, increasing the viscosity a value of η is reached for which the system (2-3) could not be solved anymore. Meanwhile, in discussions during the workshop, doubt arose about the validity of the initial condition involving the direction of the velocity, given by (6b), in the non-viscous case. Therefore, we decided to replace these conditions by

$$\frac{dy}{ds}=v_y, \qquad \frac{dx}{ds}=\sqrt{1-v_y^2}, \tag{9}$$

and tried to apply the Picard-iteration for several values of v_y. It appeared that this change of the direction of the spinning line at the orifice had a stabilizing effect on the solution of the system. Moreover,

the success of the Picard-iteration was very sensitive to the choice of the angle, determined by v_y, at the orifice (see Table 3). The quotient

η/ρ	v_y	result
0.001	0	4 iterations successfull
0.01	0	unstable
	-0.1	4 iterations successfull
	-0.2	unstable
0.1	-0.1	unstable
	-0.4	1 iteration successfull
0.2	-0.45	1 iteration successfull
0.3	-0.5	1 iteration successfull
0.4	-0.5	1 iteration successfull
0.5	-0.5	1 iteration successfull
0.7	-0.5	wrong trajectory
	-0.6	1 iteration successfull
	-0.7	1 iteration successfull

TABLE 3. Influence of η/ρ and v_y on Picard iteration

0.7 is derived from the physical values of the parameters. We present the trajectories obtained after 1 Picard-iteration for the starting values $v_y = -0.5$ and $v_y = -0.7$ in Fig. 3. It was possible to apply a

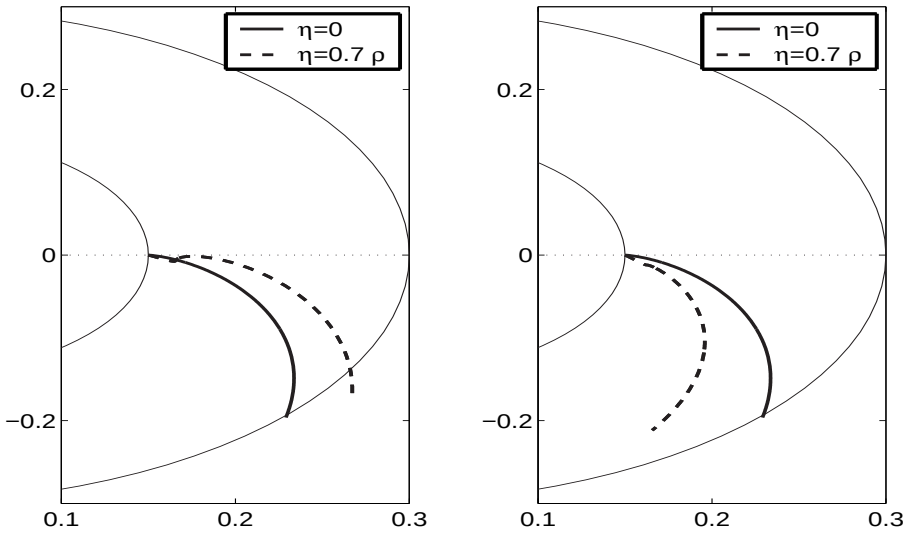


FIGURE 3. Trajectories for $v_y = -0.5$ (left) and $v_y = -0.7$ (right)

Picard-iteration in case of the initial condition $v_y = -0.5$, the resulting trajectory, however, appears to be nonphysical. We conclude that

both the solution process and the resulting trajectory are very sensitive to the choice of the initial conditions, and that some effort might be required to establish correct ones.
Secondly, we observed that the solution of the system (2-5) does not satisfy the condition (1) in the initial phase of the integration. Although (2-5) is mathematically equivalent to the original system (1-4) the numerical solutions are not, and guard must be taken that the deviation from (1) is not too strong. In a second experiment we solved the initial phase of (2-5), using the Runge-Kutta order four method with a very small step-size h, and measured the deviation from condition (1). Table 4 shows the results for several initial conditions.

F_0/Φ	$v_y = 0$	$v_y = -0.1$	$v_y = -0.2$	$v_y = -0.3$	$v_y = -0.4$	$v_y = -0.5$
0.5	$6.8_{10^{-8}}$	$6.3_{10^{-8}}$	$5.7_{10^{-7}}$	$1.6_{10^{-6}}$	$3.0_{10^{-6}}$	$4.7_{10^{-6}}$
0.7	$6.2_{10^{-7}}$	$5.8_{10^{-7}}$	$5.3_{10^{-6}}$	$1.5_{10^{-5}}$	$2.9_{10^{-5}}$	$4.7_{10^{-5}}$
0.9	$4.0_{10^{-5}}$	$3.7_{10^{-5}}$	$3.6_{10^{-4}}$	$1.0_{10^{-3}}$	$2.3_{10^{-3}}$	$4.3_{10^{-3}}$
1.1	unstable	unstable	unstable	unstable	$5.8_{10^{-2}}$	$1.1_{10^{-2}}$
1.3	unstable	unstable	$2.0_{10^{-1}}$	$1.7_{10^{-2}}$	$3.4_{10^{-3}}$	$9.2_{10^{-4}}$
1.5	$8.9_{10^{-1}}$	$3.7_{10^{-1}}$	$1.5_{10^{-2}}$	$5.1_{10^{-3}}$	$1.5_{10^{-3}}$	$5.0_{10^{-4}}$

TABLE 4. Maximum of $1 - \|[\frac{dx}{ds}, \frac{dy}{ds}]\|$ in the first 5 integration steps, $h = 10^{-5}$

Notwithstanding the small step size, condition (1) is severely violated in case v_y=0 and $F_0 > \Phi$, whereas we observe improvement for larger (negative) values for v_y. Again, we conclude that nonzero initial conditions for dy/ds should be considered. Moreover, it seems worth while to write the equations (2-3) in conservation form, using (5),

$$(10a) \quad (F - \Phi v)\frac{d^2x}{ds^2} = \frac{\Phi\omega^2}{v}\frac{dy}{ds}\left(y\frac{dx}{ds} - x\frac{dy}{ds}\right) - 2\Phi\omega\frac{dy}{ds},$$

$$(10b) \quad (F - \Phi v)\frac{d^2y}{ds^2} = -\frac{\Phi\omega^2}{v}\frac{dx}{ds}\left(y\frac{dx}{ds} - x\frac{dy}{ds}\right) + 2\Phi\omega\frac{dx}{ds},$$

and use an integration method which preserves conservation. Preliminary calculations indicate that the Runge-Kutta method does not become unstable for system (10) together with (4-5), although the accuracy is low in the case $F_0/\Phi = 1.1$. As an alternative, one might consider solving the differential algebraic system (1-4). In any case, the trajectory near the orifice is extremely sensitive to perturbations.

4. Analytical results

First we eliminate the viscous force F from the equations (2-3), and then the system is rewritten using $v = ds/dt$ to equations with time t

as independent variable, yielding

$$\begin{aligned}
\frac{d^2x}{dt^2} - \omega^2 x - 2\omega\frac{dy}{dt} &= \frac{\eta}{\rho}\frac{d}{dt}\left(\frac{1}{v^3}\frac{dv}{dt}\frac{dx}{dt}\right), \\
\frac{d^2y}{dt^2} - \omega^2 y + 2\omega\frac{dx}{dt} &= \frac{\eta}{\rho}\frac{d}{dt}\left(\frac{1}{v^3}\frac{dv}{dt}\frac{dy}{dt}\right), \\
\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 &= v^2.
\end{aligned}$$

Now, the variables x, y, v and t will be rescaled

$$\begin{aligned}
\tilde{t} &= \omega t, \quad v = \omega R_{rot}\tilde{v}, \\
x &= R_{rot}\tilde{x}, \quad y = R_{rot}\tilde{y},
\end{aligned}$$

so the equations become dimensionless. Eliminating $\tilde{v}$, dropping the $\tilde{}$ for convenience and denoting the differentiation with respect to t by $'$ yields the singularly perturbed system

$$x'' - x - 2y' = \frac{\eta}{\rho\omega R_{rot}^2}\left(\frac{(x'x'' + y'y'')\,x'}{(x'x' + y'y')^2}\right)', \tag{11a}$$

$$y'' - y + 2x' = \frac{\eta}{\rho\omega R_{rot}^2}\left(\frac{(x'x'' + y'y'')\,y'}{(x'x' + y'y')^2}\right)'. \tag{11b}$$

The initial conditions read

$$x = 1, \quad y = 0, \quad x' = \frac{v_0}{\omega R_{rot}}, \quad y' = 0,$$

and the boundary conditions at (unkown) time $t = T$

$$x(T)^2 + y(T)^2 = \frac{R_{coag}}{R_{rot}} = 4, \quad [x'(T), y'(T)] \cdot [x(T), y(T)] = 0, \cdots .$$

For the practical application we introduce the small parameters

$$\epsilon = \frac{\eta}{\rho\omega R_{rot}^2} \approx 0.12, \quad \delta = \frac{v_0}{\omega R_{rot}} \approx 0.026.$$

Then, the perturbed system reads in operator form

$$L\mathbf{u} = \epsilon\mathbf{f}(\mathbf{u}) \tag{12}$$

together with boundary conditions (BC's), and as a first step we might approximate the solution $\mathbf{u} = [x, y]$ by the regular perturbation expansion

$$\mathbf{u} = \mathbf{u_0} + \epsilon\mathbf{u_1} + \cdots ,$$

where $\mathbf{u_0}$ and $\mathbf{u_1}$ satisfy the boundary value problems

$$\begin{aligned}
L\mathbf{u_0} &= \mathbf{0}, \quad \text{(inhomogeneous BC's)}\,, \\
L\mathbf{u_1} &= \mathbf{f}(\mathbf{u_0}), \quad \text{(homogeneous BC's)}\,.
\end{aligned}$$

The solution for the first problem gives the bullet trajectory (cf. (7))

$$x_0(t) = (1+\delta t)cos(t) + tsin(t), \quad y_0(t) = -(1+\delta t)sin(t) + tcos(t),$$

The second problem is solved using Maple for various combinations of δ and ϵ. Fig. 4 shows the trajectories determined by $\mathbf{u_0}$ and $\mathbf{u_0} + \epsilon\mathbf{u_1}$ for $\delta = 0.025$ and $\epsilon = 0.0005$. The value for ϵ is non physical, but the plot clearly shows that even a small viscosity leads to an unacceptable trajectory near the rotor, probably due to the condition $y' = 0$ and to the existence of a boundary layer near the rotor.

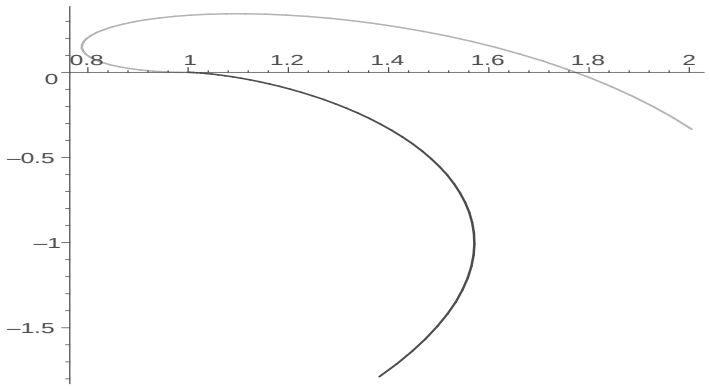


FIGURE 4. Trajectories for $\delta = 0.025, \epsilon = 0.0005$

Acceptable trajectories are obtained for values of δ not close to zero, even if the viscosity is large, as is shown in Fig. 5 (left), obtained for $\delta = \epsilon = 1$. In case of physical values for δ and ϵ it turned out to

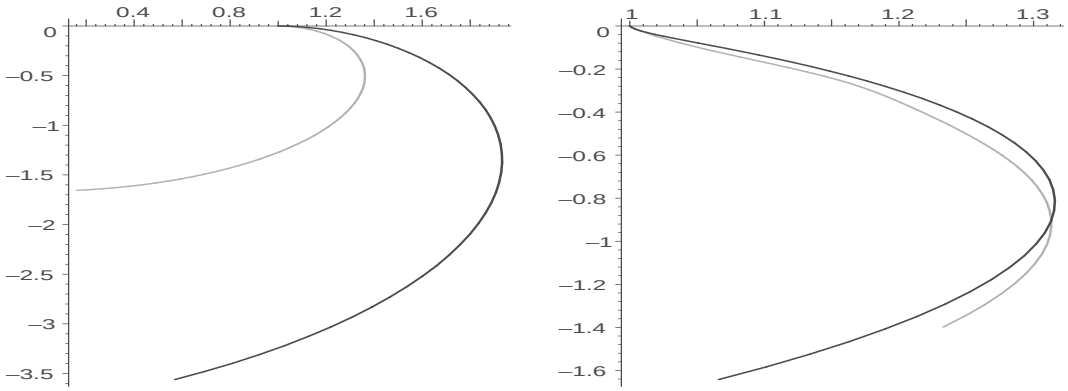


FIGURE 5. Trajectories for $\delta = 1, \epsilon = 1$ and for $\delta = 0.025, \epsilon = 0.12, y' = -sin(\pi/18)$.

be impossible to obtain satisfactory results, using $y' = 0$. However, negative angles for the trajectory at the rotor, i.e. $y' < 0$ had a stabilizing effect. In Fig. 5 (right) the results are shown for $\delta = 0.025, \epsilon =$

$0.12, y' = -sin(\pi/18)$. These results indicate again that the behavior of the solution near the rotor should be reconsidered.
System (12) can be considered as a singularly perturbed problem with a boundary layer most likely near the rotor. As a second step we could apply singular perturbation techniques in this layer. We rescale the dimensionless equations (11) near $t = 0$ by

$$x = 1 + \epsilon^{\alpha}\overline{x}, \quad y = \epsilon^{\beta}\overline{y}, \quad t = \epsilon^{\gamma}\overline{t}.$$

Interesting values for the parameters turn out to be $\alpha = \gamma = 1, \beta = 2$. Collecting the lowest order terms then gives

$$\begin{aligned} \frac{d^2\overline{x}}{d\overline{t}^2} &= \frac{d}{d\overline{t}}\left(\frac{1}{w^2}\frac{d^2\overline{x}}{d\overline{t}^2}\right), \quad w = \frac{d\overline{x}}{d\overline{t}}, \\ \frac{d^2\overline{y}}{d\overline{t}^2} + 2\frac{d\overline{x}}{d\overline{t}} &= \frac{d}{d\overline{t}}\left(\frac{1}{w^3}\frac{d^2\overline{x}}{d\overline{t}^2}\frac{d\overline{y}}{d\overline{t}}\right). \end{aligned}$$

Solving these equations will give an approximation to the trajectory in the boundary layer near the orifice. For other parts of the trajectory a similar approach, using different scalings, could be applied, and then the obtained solutions could be matched.

5. Time-dependent model

In the time-dependent description, the coordinates are r, the distance to the axis of the rotor, and t, the time. We introduce θ as the angle between the x-axis and the position vector $\boldsymbol{x}$ of the fluid,

$$\mathbf{x} = r\,(cos(\theta), sin(\theta))^T, \tag{13}$$

and we denote the velocity vector by $\mathbf{v}$. Further we introduce the *rotation matrix* J

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

so that the tangent vector and its length are given by

$$\frac{\partial \boldsymbol{x}}{\partial r} = \left(\frac{1}{r}I + \frac{\partial\theta}{\partial r}J\right)\boldsymbol{x}, \quad \left|\frac{\partial \boldsymbol{x}}{\partial r}\right| = \sqrt{1 + r^2\left(\frac{\partial\theta}{\partial r}\right)^2},$$

The normalized tangent vector $\boldsymbol{t}$ and the unit normal vector $\boldsymbol{n}$ are

$$\boldsymbol{t} = \frac{1}{|\partial\boldsymbol{x}/\partial r|}\frac{\partial\boldsymbol{x}}{\partial r}, \tag{14}$$

$$\boldsymbol{n} = J\boldsymbol{t}. \tag{15}$$

5.1. Fixed coordinate system. First we derive the time derivative of the angle θ from the fact that the fluid moves with flow velocity $\boldsymbol{v}$. The equation

$$\left(\boldsymbol{v} - \frac{\partial \boldsymbol{x}}{\partial t}\right) \cdot \boldsymbol{n} = 0,$$

then leads to, using (13) and (15),

$$\frac{\partial \theta}{\partial t} = \frac{\boldsymbol{v} \cdot \boldsymbol{n}}{\boldsymbol{x} \cdot \boldsymbol{t}}. \tag{16}$$

The other time derivatives, for A and for $\boldsymbol{v}$, will be found from the balances of mass (continuity equation) and momentum. However, we will first derive a general conservation law in a segment between the coordinates $r = R_1$ and $r = R_2$ for a local quantity d transported with flux $\boldsymbol{f}$ (see Fig. 6)

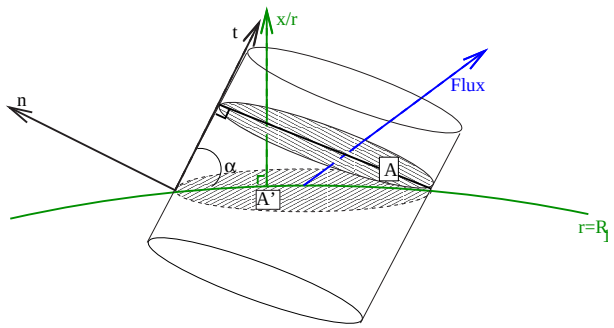


FIGURE 6. Transport passes through the 'skewed' cross-section A'.

$$\frac{\partial}{\partial t} \int_{R_1}^{R_2} Ad \left|\frac{\partial \boldsymbol{x}}{\partial r}\right| \, \mathrm{d}r + \left[\frac{\boldsymbol{x} \cdot \boldsymbol{f}}{\boldsymbol{x} \cdot \boldsymbol{t}} A\right]_{R_1}^{R_2} = 0.$$

Because this equation holds for every R_1 and R_2, we obtain the general conservation law in differential form (if the solution is sufficiently smooth):

$$\frac{\partial}{\partial t}\left(Ad\left|\frac{\partial \boldsymbol{x}}{\partial r}\right|\right) + \frac{\partial}{\partial r}\left(\frac{\boldsymbol{x} \cdot \boldsymbol{f}}{\boldsymbol{x} \cdot \boldsymbol{t}} A\right) = 0. \tag{17}$$

The continuity equation is obtained by applying (17) to mass with density $d = \rho$ and flux $\boldsymbol{f} = \rho \boldsymbol{v}$:

$$\frac{\partial}{\partial t}\left(\rho A\left|\frac{\partial \boldsymbol{x}}{\partial r}\right|\right) + \frac{\partial}{\partial r}\left(\rho A \frac{\boldsymbol{x} \cdot \boldsymbol{v}}{\boldsymbol{x} \cdot \boldsymbol{t}}\right) = 0 \tag{18}$$

The momentum equations for the x and y components of the velocity are found by applying (17) to momentum, e.g. for the x-component

$$d = \rho u, \qquad \boldsymbol{f} = \rho u \boldsymbol{v} - (1,0) T_{visc},$$

where the viscosity tensor is given by

$$T_{visc} = \frac{\eta}{|\partial \boldsymbol{x}/\partial r|} \left(\boldsymbol{t} \cdot \frac{\partial \boldsymbol{v}}{\partial r} \right) \boldsymbol{t}\boldsymbol{t}^T.$$

These equations read in vector form

$$\frac{\partial}{\partial t} \left(\rho A \boldsymbol{v} \left| \frac{\partial \boldsymbol{x}}{\partial r} \right| \right) + \frac{\partial}{\partial r} \left(\rho A \boldsymbol{v} \frac{\boldsymbol{x} \cdot \boldsymbol{v}}{\boldsymbol{x} \cdot \boldsymbol{t}} - \frac{\eta A}{|\partial \boldsymbol{x}/\partial r|} \left(\boldsymbol{t} \cdot \frac{\partial \boldsymbol{v}}{\partial r} \right) \boldsymbol{t} \right) = 0. \tag{19}$$

An equation for the kinetic energy,

$$E = \frac{1}{2} \rho A |\boldsymbol{v}|^2,$$

may be derived from

$$\frac{\partial}{\partial t} \left(\left| \frac{\partial \boldsymbol{x}}{\partial r} \right| E \right) = \boldsymbol{v} \cdot \frac{\partial}{\partial t} \left(\rho A \left| \frac{\partial \boldsymbol{x}}{\partial r} \right| \boldsymbol{v} \right) - \frac{1}{2} |\boldsymbol{v}|^2 \frac{\partial}{\partial t} \left(\rho A \left| \frac{\partial \boldsymbol{x}}{\partial r} \right| \right).$$

Substituting the continuity equation (18) and momentum equations (19), and combining advection and viscous terms yields

$$\frac{\partial}{\partial t} \left(\left| \frac{\partial \boldsymbol{x}}{\partial r} \right| E \right) + \frac{\partial}{\partial r} \left(E \frac{\boldsymbol{x} \cdot \boldsymbol{v}}{\boldsymbol{x} \cdot \boldsymbol{t}} - \frac{1}{2} \frac{\eta A}{|\partial \boldsymbol{x}/\partial r|} \boldsymbol{t}^T \frac{\partial \boldsymbol{v}\boldsymbol{v}^T}{\partial r} \boldsymbol{t} \right) = - \frac{\eta A}{|\partial \boldsymbol{x}/\partial r|} \left(\boldsymbol{t} \cdot \frac{\partial \boldsymbol{v}}{\partial r} \right)^2. \tag{20}$$

5.2. Rotating coordinate system. Stationary solutions can only be found on a rotating coordinate system. The transformation is obtained by choosing the variable $\tilde{\theta} := \theta - \omega t$, where ω is the angular speed of the rotor. In the derivation of the transformed equations we will frequently use the rotation matrix

$$C = \begin{pmatrix} \cos(\omega t) & -\sin(\omega t) \\ \sin(\omega t) & \cos(\omega t) \end{pmatrix}$$

The transformed position vector $\tilde{\boldsymbol{x}}$ is given by

$$\tilde{\boldsymbol{x}} = r \left(\cos(\tilde{\theta}), \sin(\tilde{\theta}) \right)^T, \tag{21}$$

which relates to the original coordinates by

$$\boldsymbol{x} = C \tilde{\boldsymbol{x}}.$$

The partial derivatives of $\boldsymbol{x}$ satisfy

$$\frac{\partial \boldsymbol{x}}{\partial r} = C \frac{\partial \tilde{\boldsymbol{x}}}{\partial r}, \quad \frac{\partial \boldsymbol{x}}{\partial t} = C \left(\frac{\partial \tilde{\boldsymbol{x}}}{\partial t} + \omega J \tilde{\boldsymbol{x}} \right).$$

The velocity $\boldsymbol{v}$ and the unit tangent and normal vectors can be expressed in the transformed coordinates, too:

$$\boldsymbol{t} = C\tilde{\boldsymbol{t}}, \quad \boldsymbol{n} = C\tilde{\boldsymbol{n}}, \quad \boldsymbol{v} = C\left(\tilde{\boldsymbol{v}} + \omega J\tilde{\boldsymbol{x}}\right).$$

As a consequence, the dot products satisfy

$$\boldsymbol{x}\cdot\boldsymbol{v} = \tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{v}}, \quad \boldsymbol{x}\cdot\boldsymbol{t} = \tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}, \quad \boldsymbol{t}\cdot\frac{\partial\boldsymbol{v}}{\partial r} = \tilde{\boldsymbol{t}}\cdot\frac{\partial\tilde{\boldsymbol{v}}}{\partial r}.$$

The kinematic equation now reads in the new coordinate system, using the fact that both J and C are orthogonal matrices,

$$\frac{\partial\tilde{\theta}}{\partial t} + \omega = \frac{(C\tilde{\boldsymbol{v}} + \omega CJ\tilde{\boldsymbol{x}})\cdot(C\tilde{\boldsymbol{n}})}{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}} = \frac{\tilde{\boldsymbol{v}}\cdot\tilde{\boldsymbol{n}}}{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}} + \omega. \tag{22}$$

The transformed continuity equation becomes

$$\frac{\partial}{\partial t}\left(\rho A\left|\frac{\partial\tilde{\boldsymbol{x}}}{\partial r}\right|\right) + \frac{\partial}{\partial r}\left(\rho A\frac{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{v}}}{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}}\right) = 0. \tag{23}$$

The derivation of the transformed momentum equation is slightly more complicated. First, we express each of the terms in (19) in the transformed variables.

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\rho A\boldsymbol{v}\left|\frac{\partial\boldsymbol{x}}{\partial r}\right|\right) &= C\frac{\partial}{\partial t}\left(\rho A\tilde{\boldsymbol{v}}\left|\frac{\partial\tilde{\boldsymbol{x}}}{\partial r}\right|\right) + \omega CJ\tilde{\boldsymbol{x}}\frac{\partial}{\partial t}\left(\rho A\left|\frac{\partial\tilde{\boldsymbol{x}}}{\partial r}\right|\right) + \\
&\quad + \omega CJ\rho A\left|\frac{\partial\tilde{\boldsymbol{x}}}{\partial r}\right|\frac{\partial\tilde{\boldsymbol{x}}}{\partial t} + \rho A\left|\frac{\partial\tilde{\boldsymbol{x}}}{\partial r}\right|\frac{\partial C}{\partial t}\left(\tilde{\boldsymbol{v}} + \omega J\tilde{\boldsymbol{x}}\right), \\
\frac{\partial}{\partial r}\left(\rho A\boldsymbol{v}\frac{\boldsymbol{x}\cdot\boldsymbol{v}}{\boldsymbol{x}\cdot\boldsymbol{t}}\right) &= C\frac{\partial}{\partial r}\left(\rho A\tilde{\boldsymbol{v}}\frac{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{v}}}{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}}\right) + \omega CJ\tilde{\boldsymbol{x}}\frac{\partial}{\partial r}\left(\rho A\frac{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{v}}}{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}}\right) + \\
&\quad + \omega CJ\rho A\frac{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{v}}}{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}}\frac{\partial\tilde{\boldsymbol{x}}}{\partial r},
\end{aligned}$$

and

$$-\frac{\partial}{\partial r}\left(\frac{\eta A}{|\partial\boldsymbol{x}/\partial r|}\left(\boldsymbol{t}\cdot\frac{\partial\boldsymbol{v}}{\partial r}\right)\boldsymbol{t}\right) = -C\frac{\partial}{\partial r}\left(\frac{\eta A}{|\partial\tilde{\boldsymbol{x}}/\partial r|}\left(\tilde{\boldsymbol{t}}\cdot\frac{\partial\tilde{\boldsymbol{v}}}{\partial r}\right)\tilde{\boldsymbol{t}}\right).$$

Combining these terms and using the transformed continuity equation leads to

$$\begin{aligned}
C\ &\frac{\partial}{\partial t}\left(\rho A\tilde{\boldsymbol{v}}\left|\frac{\partial\tilde{\boldsymbol{x}}}{\partial r}\right|\right) + C\frac{\partial}{\partial r}\left(\rho A\tilde{\boldsymbol{v}}\frac{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{v}}}{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}}\right) - C\frac{\partial}{\partial r}\left(\frac{\eta A}{|\partial\boldsymbol{x}/\partial r|}\left(\tilde{\boldsymbol{t}}\cdot\frac{\partial\tilde{\boldsymbol{v}}}{\partial r}\right)\tilde{\boldsymbol{t}}\right) + \\
&+\rho A\left|\frac{\partial\tilde{\boldsymbol{x}}}{\partial r}\right|\frac{\partial C}{\partial t}\left(\tilde{\boldsymbol{v}} + \omega J\tilde{\boldsymbol{x}}\right) + \omega CJ\rho A\left|\frac{\partial\tilde{\boldsymbol{x}}}{\partial r}\right|\frac{\partial\tilde{\boldsymbol{x}}}{\partial t} + \omega CJ\rho A\frac{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{v}}}{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}}\frac{\partial\tilde{\boldsymbol{x}}}{\partial r} = 0.
\end{aligned}$$

Now, observe that

$$\frac{\partial\boldsymbol{C}}{\partial t} = \omega CJ, \quad \frac{\partial\tilde{\boldsymbol{x}}}{\partial t} = J\tilde{\boldsymbol{x}}\frac{\partial\tilde{\theta}}{\partial t}.$$

Rearranging terms, and pre-multiplication by C^{-1} then yields

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\rho A\tilde{\boldsymbol{v}}\left|\frac{\partial\tilde{\boldsymbol{x}}}{\partial r}\right|\right)+\frac{\partial}{\partial r}\left(\rho A\tilde{\boldsymbol{v}}\frac{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{v}}}{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}}-\frac{\eta A}{|\partial\boldsymbol{x}/\partial r|}\left(\tilde{\boldsymbol{t}}\cdot\frac{\partial\tilde{\boldsymbol{v}}}{\partial r}\right)\tilde{\boldsymbol{t}}\right)\\
&\quad= -\rho A\left|\frac{\partial\tilde{\boldsymbol{x}}}{\partial r}\right|\omega\left(J\tilde{\boldsymbol{v}}+\omega J^2\tilde{\boldsymbol{x}}+J^2\tilde{\boldsymbol{x}}\frac{\tilde{\boldsymbol{v}}\cdot\tilde{\boldsymbol{n}}}{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}}+\frac{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{v}}}{\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}}J\tilde{\boldsymbol{t}}\right)\\
&\quad= -\rho A\left|\frac{\partial\tilde{\boldsymbol{x}}}{\partial r}\right|\omega\left(2J\tilde{\boldsymbol{v}}-\omega\tilde{\boldsymbol{x}}\right),
\end{aligned}$$

where we used (22), and the equalities

$$J^2=-I,\quad \tilde{\boldsymbol{n}}=J\tilde{\boldsymbol{t}},\quad \left(\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{t}}\right)J\tilde{\boldsymbol{v}}=\left(\tilde{\boldsymbol{x}}\cdot\tilde{\boldsymbol{v}}\right)J\tilde{\boldsymbol{t}}-\left(\tilde{\boldsymbol{v}}\cdot J\tilde{\boldsymbol{t}}\right)\tilde{\boldsymbol{x}}.$$

6. Conclusions

The viability might be questioned of the assumption that the solutions behave smoothly near the rotor. The initial condition $v_y = 0$ did lead to severe problems and non physical trajectories, both in the numerical experiments and in the regular perturbation. It has been suggested to first consider the problem of a polymer dropping down from a horizontal plate on a conveyer belt. Then, due to viscosity, the initial angle will not be perpendicular to the plate. The correct behavior near the plate might give a clue for the formulation of appropriate conditions in the rotor spinning problem.

Once correct boundary behavior has been obtained, it might very well be possible to solve the problem by clever shooting. However, care should be taken in the initial part of the trajectory, which is very sensitive to perturbations. Therefore we think that automatic shooting will fail, but trial and error by hand might be successful. In future research the models could be extended by including temperature effects and forces due to air friction.

Acknowledgement. We would like to thank Marc Peletier for his valuable contribution to our discussions.

Will the ringing of the Bourdon bell damage the Old Church Delft?

Kees Lemmens, Andrei Abramyan, Jelle Hijmissen, Hai Xiang Lin, Kees Oosterlee, Heueltje Rijnks

‡ TU Delft, EWI (DIAM), Mekelweg 4, 2628 CD Delft
C.W.J.Lemmens@ewi.tudelft.nl,
andabr33@yahoo.co.uk, J.W.Hijmissen@EWI.tudelft.nl,
H.X.Lin@ewi.tudelft.nl, C.W.Oosterlee@ewi.tudelft.nl,
H.Rijnks@ewi.tudelft.nl

1. Introduction

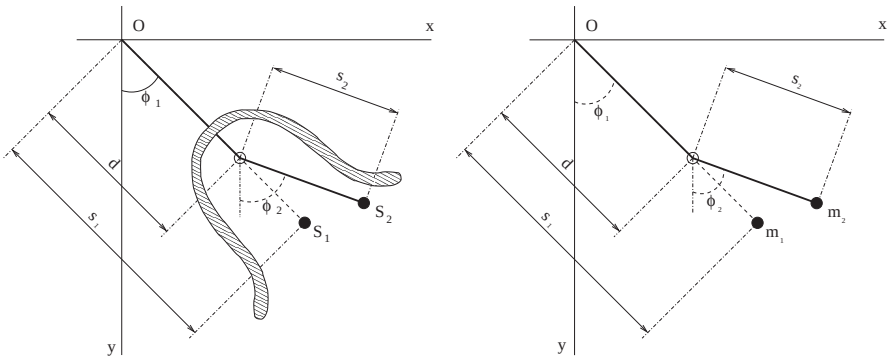
The first parish church of Delft, the old church, was built around 1200. In front of the church a 75 meters high tower, with brickwork spire and four turrets, was built in 1350. Even during its construction, the tower was plagued by subsidence. This could be because the water in the Oude Delft had to be redirected to make way for the existing church. The tower therefore was probably built on a filled-in canal. Throughout the ages, the leaning tower has been the cause of considerable alarm to many an inhabitant. The tower leans 1.20 meters to the west and 1 meter to the north.

Two unique bells hang from a heavy oak bell cage in the fourth loft in the tower of the Oude Kerk (Old Church). These are the Trinitas bell dating from 1570 and the Laudate bell dating from 1719. The Trinitas bell, or Bourdon bell, is the most exceptional of the two, weighing almost nine tonnes . The Bourdon is only rung on very special occasions such as, for example, the funeral of a member of the Dutch royal family. The powerful chime of the Bourdon causes such heavy vibrations that regular use could damage the monument.

In this paper we tried to model the effect of ringing the extremely heavy bells inside the leaning tower with mathematical methods. We modelled the leaning Old Chuch by a skew vertical Euler-Bernoulli beam and we modelled the swinging of the Bourdon bell as a non-linear singular pendulum. First in section 2 we will consider the bell model and calculate the force acting on the tower due to the swinging bell. In section 3 we will consider the skew beam model and calculate the maximum displacement in both leaning directions due to the swinging bell.

2. Mathematical model for the bell

2.1. Introduction. This section describes a mathematical model for the swing of the Bourdon bell in the tower of the Old Church in Delft. Although a bell with clapper (or bob) is formally a coupled system of 2 pendulums, we will use a single physical pendulum approximation for the computation of the pendulum period because of the relatively small mass of the clapper compared with the bell itself.
Nevertheless, as the amplitude is relatively high (about ± 70 degrees) we do not use a linear approximation such as the harmonic oscillator but use elliptic integrals to describe this pendulum in a more accurate way.
Figure (1) shows the double pendulum model and notations while figure (2) shows the single pendulum approximation.



(a) Bourdon bell with bob (b) Generalized double pendulum

FIGURE 1. Bell and bob parameters

2.2. Analysis for a non-linear singular pendulum. Using Newton's second law for rotational motion on this pendulum yields as usual for the balance of moments:

$$I_0\ddot{\varphi} = -m\ g\ s \sin(\varphi) \tag{1}$$

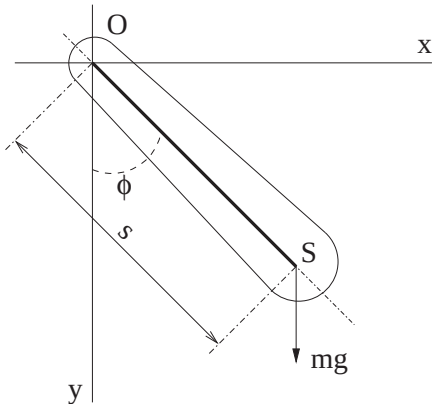
Here I_0 denotes the moment of inertia, ϕ the amplitude, m the mass, s the distance to the centre of gravity and g the gravity constant (see also fig. 2).

Substituting the commonly used "radius of gyration" i:

$$I_0 = m\ i^2 \tag{2}$$

yields:

$$\ddot{\varphi} + \frac{g\ s}{i^2} \sin(\varphi) = 0 \tag{3}$$



I_0 = moment of inertia around O

m = mass of pendulum

S = center of gravity

FIGURE 2. Single pendulum approximation (schematic)

With the angular speed ω defined by:

$$\omega = \frac{\sqrt{g\ s}}{i} \tag{4}$$

we obtain the standard form for the non-linear pendulum:

$$\ddot{\varphi} + \omega^2 \sin(\varphi) = 0 \tag{5}$$

We take as initial conditions (at $t = 0$):

$$\varphi(0) = \varphi_{max} = \alpha \text{ and } \dot{\varphi}(0) = 0$$

If we multiply (5) with $\dot{\varphi}$ and integrate we find:

$$\frac{1}{2} \int \frac{d}{dt}(\dot{\varphi})^2 = \omega^2 \int \frac{d}{dt}(\cos(\varphi + C))$$

From the initial conditions we can replace the value for C:

$$\dot{\varphi}^2 = 2\omega^2\{\cos(\varphi) - \cos(\alpha)\} \tag{6}$$

Using $1-\cos(\varphi) = 2\sin^2(\frac{\varphi}{2})$ and $1-\cos(\alpha) = 2\sin^2(\frac{\alpha}{2})$ and introducing the parameter k according to:

$$k = \sin(\frac{\alpha}{2}) \ , \ 0 \leq k \leq 1 \tag{7}$$

we can rewrite (6) into:

$$\dot{\varphi}^2 = 4k^2\omega^2(1 - \frac{\sin^2(\varphi/2)}{k^2}) \tag{8}$$

Introduction of a new variable $y = \frac{\sin(\varphi/2)}{k}$ transforms the problem into equation (10) with the help of:

$$\dot{\varphi}^2 = \frac{4k^2(\dot{y})^2}{1 - k^2y^2} \tag{9}$$

$$\dot{y} = \omega\sqrt{(1-y^2)(1-k^2y^2)} \tag{10}$$

Finally this yields as solution (11):

$$\omega\, t + C = \int_0^y \frac{d\zeta}{\sqrt{(1-\zeta^2)(1-k^2\zeta^2)}} \tag{11}$$

2.3. Elliptic integrals. The integral in the righthand side of (11) is a so-called elliptic integral of the first kind in the Legendre normal form and it's solution is available in a tabulated form ([**2**]).
It has an inverse $sn(x)$ on $0 \leq y \leq 1$, which is an elliptic function of Jacobi and where $sn(x)$ is a periodic function with a $4K$ period.
$K(k)$ is defined as follows:

$$K(k) = \int_0^1 \frac{d\eta}{\sqrt{(1-\eta^2)(1-k^2\eta^2)}} \tag{12}$$

So, on the interval $0 \leq y \leq 1$ $sn(K(k)) = k$ and $sn(K) = 1$. From (11) it follows that the solution can be written as $y = sn(\omega t + C)$. Also $sn(0) = 0$ and with $k \to 0$ we obtain $sn(x) \to \sin(x)$.
Now putting all pieces together we find:

$$y = \frac{\sin(\phi/2)}{k} = sn(\omega t + C) \tag{13}$$

With the initial conditions from above at $t = 0$ ($\phi(0) = \alpha$), we obtain

$$\sin\frac{\alpha}{2} = k\; sn(C),\ \text{or: } sn(C) = 1,\ \text{and } C = K$$

Finally, we find the solution

$$\begin{aligned} \sin\frac{\phi}{2} &= ksn(\omega\, t + K) \\ \text{and: } \phi(t) &= 2\arcsin\{sn(\omega\, t + K)\} \end{aligned} \tag{14}$$

We need the relation between $K(k)$ and k on the interval $< 0, 1 >$ before we are able to see how frequency changes with the bell amplitude. We can search a few points and present the relation between k and $K(k)$ roughly in a figure:
From (12) we find for $k = 0$:

$$K(0) = \int_0^1 \frac{d\eta}{\sqrt{1-\eta^2}} = \arcsin 1 = \pi/2$$

Also, because the integral diverges for $k \uparrow 1$ we have $\lim_{k\to 1}\ K(k) = \infty$.
For $\alpha = 90^o$ or $\pi/2$ (remember α denotes ϕ_{max}) $k = \sin\alpha/2 = 0.707$.
We can find $K(0.707)$ numerically or by looking it up in a table such as in [**2**]. Either way we obtain $K(0.707) \approx 1.854$.
These values were used to sketch fig. (3).

For a more accurate approximation we have to do some extra work:

From (14) it follows, since the period of sn is $4K$, that the period of the pendulum equals:

$$T = \frac{4K}{\omega}$$

The influence of the nonlinearity can be estimated from:

$$K(k) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}}$$

obtained by substituting $\eta = \sin\theta$ in the integral from (12).

An expansion of $(1 - k^2 \sin^2\theta)^{-1/2}$ in the binomial series gives:

$$(1 - k^2 \sin^2\theta)^{-1/2} = \sum_{n=0}^{\infty} \binom{-1/2}{n} (-k^2)^n \sin^{2n}\theta$$

Integrating term by term and using:

$$\int_0^{\pi/2} \sin^{2n}(\theta) d\theta = 2^{-2n} \binom{2n}{n} \frac{\pi}{2}$$

it follows that:

(15)

$$K(k) = \frac{\pi}{2} \sum_{n=0}^{\infty} \binom{-1/2}{n} \binom{2n}{n} (-k/2)^{2n} = \frac{\pi}{2}\{1 + \frac{k^2}{4} + \frac{9k^4}{64} + \dots\}$$

So for the period T we find:

$$T = \frac{4K}{\omega} = \frac{2\pi}{\omega}\{1 + \frac{k^2}{4} + \frac{9k^4}{64} + \dots\} \tag{16}$$

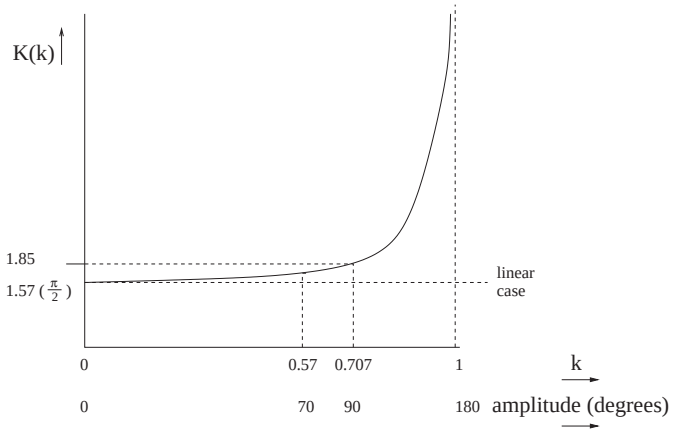


FIGURE 3. Relation between k and $K(k)$

Bell:		
Height	1820 mm	Royal Eijsbouts
Mass (without crown)	7700 kg	Royal Eijsbouts
Centre of Gravity (from bottom bell)	728 mm	Royal Eijsbouts
Moment of inertia in CG (S)	5110 $kg.m^2$	Royal Eijsbouts
Rotation axis (from bottom bell)	1460 mm	TNO Delft
Crown:		
Mass	550 kg	Royal Eijsbouts
Distance from rotation axis	300 mm	estimated
Counterweight:		
Mass	1000 kg	estimated
Distance from rotation axis	1100 mm	TNO Delft / estimated

TABLE 5. Some properties of the Bourdon bell

where the first term $2\pi/\omega$ represents the period of the linearized system (the harmonic oscillator!).

For $k = 0.707$ (or the amplitude $\alpha = \frac{\pi}{2}$), it follows by using the first 2 terms:

$$k^2/4 + 9k^4/64 \approx 0.16$$

So for an amplitude of $\frac{\pi}{2}$ the nonlinear terms raise T with 16% as compared to the linearized case.

$K(0.707) \approx \frac{\pi}{2}\,(1+0.16) \approx 1.83$ whereas 1.854 is the value from ([**2**] we already used above.

For k close to 1 (amplitude $\alpha = \pi$) the series converges only slowly and there we need more and more terms to find a reasonable accurate approximation, but for the Bourdon bell this is not necessary.

2.4. Computing the bell period. To be able to use the formulas above we need detailed values for many properties of the Bourdon bell. Unfortunately some of the values below are at this moment only rough estimations.
The Royal Eijsbouts Company and TNO Delft (both in the Netherlands) provided us with some numbers. Other were estimated by ourselves using photographs. The most important ones are listed in table (5).

If we assume that the maximum amplitude of the Bourdon bell $\alpha = 70^o (= 7/18\pi)$ we have $k = \sin(\frac{\alpha}{2}) = 0.57$. Using the formula from (15)

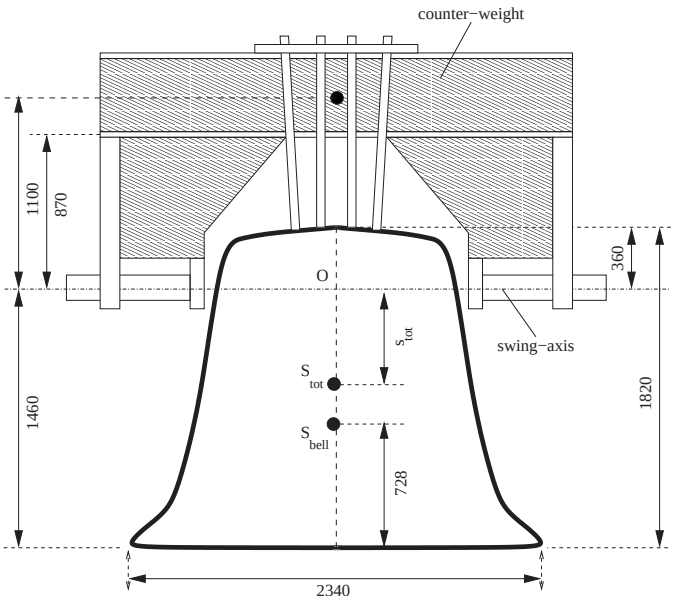


FIGURE 4. Bourdon bell dimensions

or the graph from figure (3) we find:

$$K(0.57) \approx \frac{\pi}{2}\{1 + \frac{0.57^2}{2} + 9\frac{0.57^4}{64}\} = 1.73$$

The Bourdon bell has a counterweight on the opposite side of the rotation axis O (see also figure (4)). This implies that the new center of gravity S_{tot} for the total combined system must be shifted towards O. This new center can be easily computed:

$$\begin{aligned} M_{counter}.(S_{counter} + s_{tot}) &= M_{bell}.(S_{bell} - s_{tot}) \\ 1000.(1100 + s_{tot}) &= 7700.(732 - s_{tot}) \\ s_{tot} &= 4536400/8700 = 521\ mm\ (0.52\ m) \end{aligned}$$

M_{tot} bell + crown + counterbalance: 9250 kg
I_{bell} around O (using Steiners rule): $5110 + 7700.(732)^2 = 9235\ kg.m^2$
Total moment of inertia:

$$\begin{aligned} I_0 &= I_{bell} + I_{counter} + I_{crown} \\ &= 9235 + 1000.(1.1)^2 + 550.(0.03)^2 \\ &\approx 10500\ kg.m^2 \end{aligned}$$

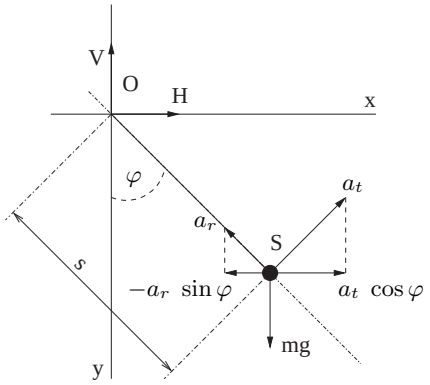


FIGURE 5. Contribution of accelerations to force (here only shown for H)

Now we can gather all other numbers and compute the period:

$$\begin{aligned} i &= \sqrt{\frac{I_0}{M}} = \sqrt{\frac{10500}{9250}} \approx 1.07\ m \\ \omega &= \frac{\sqrt{g\ s_{tot}}}{i} = \frac{\sqrt{9.81.0.52}}{1.07} \approx 2.11\ rad/sec. \\ T &= \frac{4\ K}{\omega} = \frac{4.1.73}{2.11} \approx \mathbf{3.28\ sec} \end{aligned}$$

As there are about 2 chimes per period (the clapper hits the bell twice per period) this implies $n = \frac{2}{T}.60 = \frac{2\,.\,60}{3.28} \approx 37$ chimes/min.

Linear approximation: If we would have used the linear approach (simple harmonic oscillator) the result would be $T_{lin} = \frac{2\Pi}{\omega} = 2.98$ and 40 chimes/min.
So the difference is $\frac{K(0)}{K(\alpha)} = \frac{1.57}{1.73} = 0.91$ or about 9 %.

2.5. The forces on the axis of rotation. The acceleration of the centre of gravity S consists of 2 components (see also figure 5):

$$\begin{aligned} \text{in radial direction} &: a_r = s\dot{\varphi}^2 \\ \text{in tangential direction} &: a_t = s\ddot{\varphi} \end{aligned}$$

So using the contributions of both in the x-direction we find:

$$\begin{aligned} \ddot{x}_s &= -a_r \sin\varphi + a_t \cos\varphi \\ &= -s\ \dot{\varphi}^2 \sin\varphi + s\ddot{\varphi}\cos\varphi \\ &= -s\ \omega^2 \sin\varphi\{4(k^2 - \sin^2(\frac{\varphi}{2})) + \cos\varphi\} \end{aligned}$$

The last line was obtained by using equations (5) and (8)).

With $\sum \vec{F}_x = m\vec{\ddot{x}}_s$ the horizontal force H on the pendulum in the reference point 0 (to the rightside) is given by:

$$H = -m\ s\ \omega^2\{4(k^2 - \sin^2(\frac{\varphi}{2}))\sin\varphi + \sin\varphi cos\varphi\} \tag{17}$$

and with $\sum \vec{F_y} = m\vec{\ddot{y}_s}$ the vertical force V in the reference point 0 (in upwards direction) is given in a similar way by:

$$\ddot{y}_s = a_r\cos\varphi + a_t\sin\varphi = s\ \omega^2\{4(k^2 - \sin^2(\frac{\varphi}{2}))\cos\varphi - \sin^2\varphi\}$$

This yields by using the same substitutions as for the horizontal force:

$$V = m\ s\ \omega^2\{4(k^2 - \sin^2(\frac{\varphi}{2}))\cos\varphi - \sin^2\varphi\} + m\ g \tag{18}$$

Here of course φ is defined by equation (5).

By taking the derivative for H and V and see where these are equal to 0 we obtain the maximum amplitudes for H and V. Doing this in Maple yields:

$$\begin{aligned} V_{max} \text{ occurs for } &\quad \phi = &&\ 0\ rad. \\ H_{max} \text{ occurs for } &\quad \phi \approx &&\ 0.7\ rad. \end{aligned}$$

Using the numerical values we found in the previous section this yields for the forces:

$$\begin{aligned} H &\approx 23500\ N \\ V &\approx 115000\ N \text{ (or 30300 N without the } m\ g \text{ component)} \end{aligned}$$

Note that the signs of the forces are relative to the directions as shown in figure (5).

2.6. Conclusion. Given the fact that for several parameters we only have a very rough estimation (especially for the counterweight), the error caused by a linear approximation of the system is after all not such a big issue and may be even less than the error caused by incorrect parameters ...

The bell is only tolled at very special occasions of which the most important one is a funeral of a member of the Royal Dutch family, which is of course not a very common occasion.
However, just one day after the workshop ended the former Queen of the Netherlands Juliana died unexpectedly and so the bell was tolled at her funeral only a couple of days later.
We were able to obtain a short video recording from the swinging bell at the funeral, so we could measure the period straight from this video.

This turned out to be **3.2 sec** or about 37 chimes/min, so notwithstanding the possibly incorrect parameters the results from above resemble reality quite well!

3. The forces acting on the tower

The Old Church Tower consists of 4 parts with total length 75 meters. The bell of the tower is located at the height of 44 meters.The rest upper part of the tower has a much smaller weight than the previous parts. So we decide to take into account the influence of this part on the behavior of the tower construction as an additional weight which is uniformly distributed onto other parts. Different parts of the tower have different areas of cross sections and subsequently with different moments of inertia. The beam-like model can be applied for a first rough estimation of the displacement of the construction. In shipbuilding such model is often applied for the estimation of the hull behavior under the action of its weight and for vibration calculations. Because of the leaning the forces which are acting on the tower construction depend on the angle of leaning. The weight of the tower can be modeled as a load which is distributed according to the linear dependence on the length of the tower. The weight has two projections: one is along the tower axis and another perpendicular to the axis. In the process of ringing the dynamic force appears as a result of an action of a moving bell. This force also has two projections which change its values harmonically in time. So, finally the tower is modelled as a Euler-Bernoulli beam under the action of a linearly distributed weight along the length and a dynamic force caused by the bell.

For calculation of the eigen frequencies nevertheless another model of a beam should be introduced, because the ratio of the tower length to its width is 1:4.4 and the shear and longitudinal forces should be taken into account in computations of natural frequencies.

The boundary conditions for such a model unfortunately cannot be found exactly, because there is no data concerning the properties of the tower foundation and the soil properties. The boundary conditions which were considered are: one edge is fixed and another is free. For that type of boundary conditions and for fixed distribution of external forces, the displacements of the construction will be largest from all possible real boundary conditions. It is a special task to find out what boundary conditions should be taken into account for the real tower. In the next section, the governing equation for the tower structure will be described.

3.1. Beam equation. As a first approximation we model the skew tower by a skew Euler-Bernouilli beam:

$$(EI(x)u_{xx}(x,t))_{xx} + (T(x,t)u_x(x,t))_x + \rho A u_{tt}(x,t) = q(x,t)/L, \tag{19}$$

where $EI(x)$ is the bending stiffness, ρ the density, A the area of the beam, $T(x,t)$ the longitudinal compressible force due to the acceleration due to gravity(g) and the bell dynamic force and where $q(x,t)$ is the dynamic force. The longitudinal compression force is given by $W(x) = g_x[(1-x)u_x(x,t)]_x$ where if the tower has a leaning of α degrees, $g_x = g\cos(\alpha)$ and $g_y = g\sin(\alpha)$ and where $W(x)$ is the mass of the beam along its length and $u(x,t)$ the displacement of the beam along its length.

The governing system of equations has the following form:

$$(EI_1 u_{1_{xx}}(x,t))_{xx} + (T_1(x,t)u_{1_x}(x,t))_x + \rho A_1 u_{1_{tt}}(x,t) = q_1(x,t), \tag{20}$$

$$(EI_2 u_{2_{xx}}(x,t))_{xx} + (T_2(x,t)u_{2_x}(x,t))_x + \rho A_2 u_{2_{tt}}(x,t) = q_2(x,t), \tag{21}$$

$$(EI_3 u_{3_{xx}}(x,t))_{xx} + (T_3(x,t)u_{3_x}(x,t))_x + \rho A_3 u_{3_{tt}}(x,t) = q_3(x,t), \tag{22}$$

$$u_1(0,t) = u_{1_x}(0,t) = 0, \tag{23}$$

$$u_{3_{xx}}(L,t) = u_{3_{xxx}}(L,t) = 0. \tag{24}$$

with the following conditions at $x = L_1$ and $x = L_2$:

$$u_i(L_1,t) = u_j(L_i,t), \tag{25}$$

$$u_{i_x}(L_i,t) = u_{i+1_x}(L_i,t), \tag{26}$$

$$EI_i u_{i_{xx}}(L_i,t) = EI_{i+1} u_{i+1_{xx}}(L_i,t), \tag{27}$$

$$EI_i u_{i_{xxx}}(L_i,t) = EI_{i+1} u_{i+1_{xxx}}(L_i,t). \tag{28}$$

where $i = 1$ and $i = 2$ and where EI_i, A_i, T_i and $q_i(x,t)$ for $i = 1,2,3$ are the bending stiffness, the area the longitudinal compression force and the dynamic force of the three parts of the beam respectively.

Assume that $I(x)$ is a constant, then introduce the dimensionless variables $x = \frac{x}{L}$, $u(x,t) = \frac{u(x,t)}{L}$ and $t = t\sqrt{\frac{\rho A_2 L^4}{EI}}$, to put the equations (20)-(28) in the following non-dimensional form

$$u_{i_{xxxx}}(x,t) + \epsilon_i \left(\hat{W}(x)u_{i_x}(x,t)\right)_x + u_{i_{tt}}(x,t) = \alpha\epsilon_i \hat{W}(x,t), \tag{29}$$

where $i = 1,2,3$ and where $\epsilon_i = \frac{g_x L^3 \rho A_2}{EI_i}$, $\alpha = \frac{g_y}{g_x}$ and

$$\hat{W}(x) = \begin{array}{l} \frac{3}{4} + \frac{A_1}{A_2}\left(\frac{1}{4} - x\right), 0 \le x \le \frac{1}{4}, \\ \frac{1}{4} + \left(\frac{3}{4} - x\right), \frac{1}{4} \le x \le \frac{3}{4}, \\ 1 - x, \frac{3}{4} \le x \le 1. \end{array} \tag{30}$$

3.2. The static problem. First we consider the static problem. The static equations describing the motion of the church are given by

$$u_{1_xxxx}(x) + \epsilon_1 \left(\hat{W}(x)u_{1_x}(x)\right)_x = \alpha\epsilon_1\hat{W}(x), \tag{31}$$

$$u_{2_xxxx}(x) + \epsilon_2 \left(\hat{W}(x)u_{2_x}(x)\right)_x = \alpha\epsilon_2\hat{W}(x), \tag{32}$$

$$u_{3_xxxx}(x) + \epsilon_3 \left(\hat{W}(x)u_{3_x}(x)\right)_x = \alpha\epsilon_3\hat{W}(x), \tag{33}$$

$$u_1(0) = u_{1_x}(0) = u_{1_{xx}}(1) = u_{1_{xxx}}(1) = 0, \tag{34}$$

with the following conditions at $x = L_1 = \frac{1}{4}$ and $x = L_2 = \frac{3}{4}$

$$u_i(L_i, t) = u_{i+1}(L_i, t), \tag{35}$$

$$u_{i_x}(L_i, t) = u_{i+1_x}(L_i, t), \tag{36}$$

$$EI_i u_{i_{xx}}(L_i, t) = EI_{i+1}u_{i+1_{xx}}(L_i, t), \tag{37}$$

$$EI_i u_{i_{xxx}}(L_i, t) = EI_{i+1}u_{i+1_{xxx}}(L_i, t). \tag{38}$$

where $i = 1$ and $i = 2$.
If we have the values of ϵ_i and α we can determine the maximum static displacement of the beam. The tower is leaning in two directions (y and z). The first leaning is $2.5m$ and the second leaning is $0.5m$. So

$$\begin{aligned} \alpha_y &= \frac{g_x}{g_y} = \frac{2.5}{44} = 0.056, \\ \alpha_z &= \frac{g_x}{g_y} = \frac{0.5}{44} = 0.011. \end{aligned}$$

If the mass of the tower is $10^7 kg$ and the density is for every part is the same. The value for the product $\rho A_2 L$ for the tower is $0.94176 * 10^7 kg$. For the other properties we have

$$E = 2.5 * 10^9 Pa, \tag{39}$$

$$L = 44m, \tag{40}$$

$$g_x = 9.8m/s^2, \tag{41}$$

$$I_1 = 1900m^4, \tag{42}$$

$$I_2 = 1510m^4, \tag{43}$$

$$I_3 = 1080m^4. \tag{44}$$

Then we have $\epsilon_1 = 0.0376$, $\epsilon_2 = 0.0473$ and $\epsilon_3 = 0.0655$. The maximum displacements in the static case in y-direction and z-direction are given by

$$\begin{aligned} u_{max,static,y} &= 3.3mm, \\ u_{max,static,z} &= 0.67mm, \end{aligned}$$

respectively, where the subscript y, z denote the direction of the leaning. So the maximum displacement t of the tower is

$$u_{max,static} = \sqrt{u^2_{max,static,y} + u^2_{max,static,z}} = 3.4mm. \tag{45}$$

3.3. The dynamic problem. Now we consider the dynamic problemand so we also include the dynamic forces due to the swinging of the bell. We have the following partial differential equation for the ith part$(i = 1, 2, 3)$ of the beam

$$\begin{aligned} u_{i_x xxx}(x,t) + \epsilon_i \Big((\hat{W}(x) &+ B_1 \cos(\alpha)\cos(\omega t)) u_{i_x}(x,t) \big)_x \\ &= \alpha \epsilon_i \hat{W}(x) + B_2 \sin(\alpha)\cos(\omega t) \end{aligned} \tag{46}$$

where $B_1 \cos(\alpha) = \frac{\hat{B}_1}{g\rho AL} = 0.0010$, $B_2 \sin(\alpha) = \frac{\hat{B}_2 L^2}{EI} = 5.1 * 10^{-5}$ and $\omega = \hat{\omega}\sqrt{\frac{\rho A_2 L^4}{EI}} = 0.50$. Here we used that $\hat{B}_1 = -1.1 * 10^5 N$, $\hat{B}_2 = -7.0 * 10^3 N$ and $\hat{\omega} = 2.18$.

Now we will consider the dynamic problem. For the dynamic problem we consider the three parts of the beam to have the same moment of inertia and area, the moment of inertia of the second part(i.e. $I = 1510m^4$). It is possible to use this very rough estimation. We suppose $u(x,0) = u_t(x,0) = 0$. So, we consider the following initial-boundary value problem describing the oscillations of the Old Churh due to the swiging of the bell:

$$\begin{aligned} u_{xxxx}(x,t) + \epsilon \big((\gamma + (1-x)) + \epsilon^2 b_1) u_x(x,t) \big)_x &+ u_{tt}(x,t) \\ &= \hat{\alpha}\epsilon^2(1-x) + \epsilon^4 b_2 \cos(\omega t), \end{aligned} \tag{47}$$

$$u_(0,t) = u_x(0,t) = u_{xx}(1,t) = 0, \tag{48}$$

$$\epsilon^2 \delta u_x(1,t) + \delta\epsilon u_{tt} - u_{xxx}(1,t) = 0, \tag{49}$$

$$u(x,0) = u_t(x,0) = 0. \tag{50}$$

The value for ϵ we considered in the previous section and is given by $\epsilon_2 = 0.047$. So, other parameters can be written as a product of ϵ. In this way the parameters b_1, b_2 and $\hat{\alpha}$ can be calculated and are given by $\delta = 1.6$, $b_1 = 0.45$, $b_2 = 1.0$, $\hat{\alpha}_1 = 1.2$ and $\hat{\alpha}_2 = 0.23$.

We will use a two-time scale perturbation method to solve this problem. The perturbation method is used because the equation (47) can not be solved analytical. In this method the solution is supposed to be a series of the eigenfunctions of the unperturbated problem(i.e. problem (47)-(50) with $\epsilon = 0$). We will only truncate to the first three eigenfunctions to construct an $O(\epsilon^3)$-approximation. Using this method we derive the following values

$$\begin{aligned} u_{max,dynamic,y} &= 12mm, \\ u_{max,dynamic,z} &= 2.2mm, \end{aligned}$$

for the maximum displacement of the tower in the y and z-direction respectively. So the maximum displacement of the tower is

$$u_{max,dynamic} = \sqrt{u^2_{max,dynamic,y} + u^2_{max,dynamic,z}} \cong 12mm. \tag{51}$$

This maximum displacement is at the top of the beam.

4. Conclusion

In this paper we considered the maximum dynamic displacement of the Old Church in Delft due to the swinging of the bell, the so-called Bourdon. We modelled the leaning Old Chuch by a skew vetical Euler-Bernoulli beam and we modelled the swinging of the Bourdon bell as a non-linear singular pendulum. First we obtain the values of the forces due to the bell acting on the Churh. We used these values to obtain that the maximum dynamic displacement of the Churh is $12mm$.

References

[1] Ferdinand Verhulst, *Nonlinear Differential Equations and Dynamical Systems*, (2th ed.), (Berlin: Springer Verlag, 1996).

[2] Jahnke-Ende, *Tafeln Höhern Funktionen*, (5th ed.), (Leipzig: B.G. Teubner, 1960).

[3] Maria L. Beconcini, Stefano Bennati, Walter Salvatore, *Structural characterisation of a medieval bell tower: First historical, experimental and numerical investigations*, (University of Pisa, Dept. of structural engineering, Pisa, Italy, 2001).

Environmental effects of traffic

Peter Sonneveld

TU Delft, EWI (DIAM), Mekelweg 4, 2628 CD Delft, p.sonneveld@ewi.tudelft.nl

1. Problem formulation

Roads, highways etc cause noise, air pollution and other unwished contamination of the environment.
For a road system these effects can be modelled in terms of regions (buffer zones), defined for parts of the road called 'sections', in which the negative effects are considered to be unsustainable high for combination with human inhabitance (for example).
In the design process of (parts of) a road, these effects are dealt with by determining its environmental load to living areas. This quantity is identified with the surface area of the intersection of the buffer zones with the actual domain where people live. For the latter, postal code zones can be used.
The problem handled by the DEMIS group is

> Development of an algorithm, for determining the environmental load of each section of the 'road in design' in the design phase.

The following restrictions and simplifications are to be used in the algorithm:

(1) The algorithm must be implementable as interactive tool, in order to adapt the road design for decreasing the environmental load.
(2) Buffer zones are defined as polygons, and so are the postal code zones.
(3) Overlapping parts of road segments must be counted properly. In the present modelling this means 'counted only once'
(4) Therefore one has to deal with handling the overlap between the buffer zones of different sections. A (possible) complication: *Overlapping buffer zones do not necessarily belong to adjacent road segments.*

2. Participants of the work group

The following people have been working on this problem

– Poul Grashoff (*DEMIS, Initiator*)

– Ewa Matusiak (*Un of Vienna*)
– Etelvina Perez (*TU Delft*)
– Fahmi Naifar (*TU Delft*)
– Bob Planqué (*CWI*)
– Valeriu Savcenco (*CWI*)
– Peter Sonneveld (*TU Delft*)

3. Two approaches

After a first meeting of the complete group, two different approaches appeared to be promising.

1. Road segment centered approach. (*Bob, Valeriu & Fahmi*) This approach is based on the idea originated at DEMIS:
Make a description of the complete road in terms of non-overlapping polygons. The, from mathematical point of view rather simple, operations for doing this can be done using polygon manipulation tools that can be found in Matlab. (Allbeit only after some non-trivial search procedure)
At first, there was some disagreement on which parts of the intersections of more road-sections, should "belong to" which road section. In the case of a complicated road crossing situation, a cumulated overlap of many buffer zone polygons may require addition of environmental loads, at least for the air pollution component.
However, taking into account the simplicity of the environmental model, the group working on this idea decided to assign the mutual overlap to the first polygon in the system it is part from, assuming that the polygons are ordered some way. Proceeding like this, a purely additive construction can be chosen, which is of importance for the adaptivity of the algorithm.
The obvious question about the cumulative effects of more road sections on one postal code zone, could not be implemented, since also these effects are beyond the scope of the model.

2. Postal code zone centered approach. (*Eva & Etelvina*) In this alternative approach, the basis is the set of postal code zones. The positioning of the complete road is determined by the shortest path of which the buffer zone doesn't intersect any postal code zone. If the road must satisfy other types of restriction (which is always the case) then these can be met by penalty functions (for instance virtual postal code areas).
One advantage of this approach can be the possibility of actual minimization of the costs. One possible drawback can be that the resulting optimal road is far away from optimal in an other respect, like for instance an excessive length. This could be solved by modifying the concept "length."

The required software for this approach can be based on software that is standard in the world of optimization. The computational complexity can be reduced by using initially simple polygonal hulls for postal code zones with complicated shape. Fine tuning can be done by a mix of this approach and the road segment centered approach.
At the end of the *MFI*-week, this approach was only conceptually worked out.

4. First approach

Overview of the activities. After having chosen approach 1, the group decided quickly to use Matlab as programming platform. Some of the reasons:

(1) Every one was familiar with this environment.
(2) It is present in this university and on the laptop of one of the participants.
(3) For large multidimensional problems, it is an nearly optimal environment. So if someone wants to develop the required software from scratch in C or Fortran, he probably can't beat a well designed Matlab code, and certainly not in a finite time.
(4) If the software is to be used in a non-Matlab enviromnment, a stand-alone version of the software can be made with help of Matlab. Although none in our group has experience in that field, this seems to be not difficult.

Software for calculating intersections or unions of polygonal areas is basically simple. But practically not quite so simple. Luckily there exist free polygon clipping software for Matlab. In using this, the group met a first problem: The software was not fully grown up.
Apparently convex regions were expected, and some intersections couldn't be expressed at all, because they were non-connected, or multi-connected.
As a matter of fact, the activities of the 'approach 2 - group' were considerably frustrated by these circumstances.
After some search on the web and in the personal network of one of the group members, a version of the software was found that seemed to be stable and reliable.
The mathematical set up for approach 1 can be formulated simply as follows. Denote the polygonal representation of the bufferzones by $B_1, \ldots, B_N$. Now define the sequences of domains $\{R_1, \ldots, R_N\}$ and $\{C_1, \ldots C_N\}$ by

$$R_1 = D_1, \quad C_k = D_k \setminus R_{k-1}, \quad R_k = R_{k-1} \cup C_k,\ k = 2, 3, \ldots, N$$

The sequence $\{C_k\}$ consists of disjoint regions, and can be used for calculating the environmental load.

At the presentation, the group could show some elementary results on the set up of the program, and some actual intersection examples based on actual data of the road system around Rotterdam.

Epilog. At first sight the zoning problem seems a rather simple straightforward problem. The difficulties are more in the modelling of the real problem, than in the actual mathematics of the chosen model. In fact the participants of the 'approach 1-group' had a strong feeling that it was a mere programming job. And if there hadn't been free polygon-clipping-software, I would have started right away writing it myself.
Now there *was* that type of software, so thinking stopped, and the work reduced to pushing buttons until the vehicle would be under control.
Some time after the workshop, I still hadn't received a fault-free version of the clippol package, nor I succeeded in downloading it myself. So I started to write my own. And discovered that this kind of manipulation isn't trivial at all! Using rectangle-like regions, building the required union leads to terrible kinds of non-connected and multi-connected regions. Not difficult of course, but it requires a suitable datastructure to handle that, and, more important, some thinking.
Apparently the other members of the group experienced similar things, and so the project felt asleep.
The question now is: did we achieve nothing at all? I think that 's not the case

Bad news: My own later experiments showed that a proper additive procedure is not possible. If a complete assemblage of bufferzones has been carried out, and bufferzone D_k is modified, in order to decrease the environmental load, the overlapping parts with bufferzones D_j, $j > k$ become active. But these themselves may consist partly of other overlapping zones. Keeping an administration of the overlaps is really a nasty job.
This we could have seen the first afternoon, if we had thought a little longer.

Good news: The complete calculation of the environmental load for a trial design of a road system can probably be done in a reasonably short time, as far as I may consider the scale of my experiments as representative for the real problem. In that case a completely new calculation could be used as a 'response' to any modification in an interactive process. This is of course a brute force solution, but what's wrong with that?

Final conclusion. In my opinion, this interesting problem deserved some more analysis time than we gave to it. This is due to the combination of the short period we had to show some results, the quick

recognition of the fact it was a programming job, and the fact that we under-estimated the complications of precisely *that* part of the problem. If we had decided to start with making a mathematical analysis, as should be done *always*, there wouldn't have been a working program on a lab-scale, showing a piece of the requested work. But a reliable outline on how to handle the problem, and why it should work, could have been finished.

Finally, I think that usually an element of competition is present in work shops like this, and competition is not always improving the results of scientific activities, at least not in a contest of one week.