A Note on Adjoint Linear Algebra

Uwe Naumann[1[0000−0002−7518−5922]]

Department of Computer Science, RWTH Aachen University, 52056 Aachen,
Germany naumann@stce.rwth-aachen.de
www.stce.rwth-aachen.de

Abstract. A new proof for adjoint systems of linear equations is presented. The argument is built on the principles of Algorithmic Differentiation. Application to scalar multiplication sets the base line. Generalization yields adjoint inner vector, matrix-vector, and matrix-matrix products leading to an alternative proof for first- as well as higher-order adjoint linear systems.

Keywords: algorithmic differentiation · adjoint · linear algebra.

1 Motivation

Algorithmic Differentiation [3, 5] of numerical programs builds on a set of elemental functions with known partial derivatives with respect to their arguments at the given point of evaluation. The propagation of adjoint derivatives relies on the associativity of the chain rule of differential calculus. Differentiable combinations of elemental functions yield higher-level elementals. Efficient implementation of AD requires the highest possible level of elemental functions.

Basic AD assumes the set of elemental functions to be formed by the arithmetic operators and intrinsic functions built into the given programming language. While its application to linear algebra methods turns out to be straight forward basic AD is certainly not the method of choice from the point of view of computational efficiency. Elementals of the highest possible level should be used. Their derivatives should be formulated as functions of high-level elementals in order to exploit benefits of corresponding optimized implementations.

Following this rationale this note presents a new way to derive adjoint systems of linear equations based on adjoint Basic Linear Algebra Subprograms (BLAS) [4]. It is well known (see [2] and references therein) that for systems $A \cdot \mathbf{x} = \mathbf{b}$ of n linear equations with invertible A and *primal* solution $\mathbf{x} = A^{-1} \cdot \mathbf{b}$ first-order *adjoints* $A_{(1)}$ of A (both $\in \mathbb{R}^{n\times n}$ with $\mathbb{R}$ denoting the real numbers) and $\mathbf{b}_{(1)}$ of $\mathbf{b}$ (both $\in \mathbb{R}^n$) can be evaluated at the primal solution $\mathbf{x} \in \mathbb{R}^n$ as

$$\begin{pmatrix} \mathbf{b}_{(1)} = A^{-T} \cdot \mathbf{x}_{(1)} \\ A_{(1)} = -\mathbf{b}_{(1)} \cdot \mathbf{x}^T \end{pmatrix} . \tag{1}$$

The main contribution of this note is an alternative proof for Eqn. (1) that builds naturally on the adjoint BLAS used in the context of state of the art AD. For

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08754-7_27

consistency with related work we follow the notation in [5], that is, $v^{(1)} \in \mathbb{R}$ denotes the value of the first-order directional derivative (or tangent) associated with a variable $v \in \mathbb{R}$ and $v_{(1)} \in \mathbb{R}$ denotes the value of its adjoint.

2 Prerequisites

The Jacobian $\nabla F = \nabla F(\mathbf{x}) \equiv \frac{dF}{d\mathbf{x}}(\mathbf{x}) \in \mathbb{R}^{m \times n}$ of a differentiable implementation of $\mathbf{y} = F(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$ as a computer program induces a linear mapping $\mathbf{y}^{(1)} = \nabla F \cdot \mathbf{x}^{(1)} : \mathbb{R}^n \to \mathbb{R}^m$ implementing the tangent of F. The corresponding *adjoint* operator $\nabla F^* = \nabla F^*(\mathbf{x})$ is formally defined via the inner vector product identity

$$\langle \nabla F \cdot \mathbf{x}^{(1)}, \mathbf{y}_{(1)} \rangle = \langle \mathbf{x}^{(1)}, \nabla F^* \cdot \mathbf{y}_{(1)} \rangle \tag{2}$$

yielding $\nabla F^* = \nabla F^T$ [1]. In the following all (program) variables are assumed to be alias- and context-free, that is, distinct variables do not overlap in memory and F is assumed to be not embedded in an enclosing computation. We distinguish between *active* and *passive* variables. Derivatives of all active outputs of the given program are computed with respect to all active inputs. We are not interested in derivatives of passive outputs nor are we computing derivatives with respect to passive inputs.

3 BLAS Revisited

In its basic form AD builds on known tangents and adjoints of the arithmetic functions and operators built into programming languages. Tangents and adjoints are propagated along the flow of data according to the chain rule of differential calculus. We enumerate entries of vectors $\mathbf{v} \in \mathbb{R}^n$ staring from zero as $v_0, \ldots, v_{n-1}$.

From the perspective of AD adjoint versions of higher-level BLAS are derived as adjoints of lower-level BLAS. Optimization of the result aims for implementation using the highest possible level of BLAS. For example, adjoint matrix-matrix multiplication (level-3 BLAS) is derived from adjoint matrix-vector multiplication (level-2 BLAS) yielding efficient evaluation as two matrix-matrix products (level-3 BLAS) as shown in Lemma 4. Rigorous derivation of this result requires bottom-up investigation of the BLAS hierarchy. We start with basic scalar multiplication (Lemma 1) followed by the inner vector (Lemma 2) and matrix-vector (Lemma 3) products as prerequisites for the matrix-matrix product.

Lemma 1. *The adjoint of scalar multiplication $y = a \cdot x$ with active $a, x, y \in \mathbb{R}$ is computed as*

$$\begin{aligned} a_{(1)} &= x \cdot y_{(1)} \\ x_{(1)} &= a \cdot y_{(1)} \end{aligned} \tag{3}$$

for $y_{(1)} \in \mathbb{R}$ yielding $a_{(1)}, x_{(1)} \in \mathbb{R}$.

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08754-7_27

Proof. Differentiation of $y = a \cdot x$ with respect to a and x yields the tangent

$$y^{(1)} = \left\langle \begin{pmatrix} a^{(1)} \\ x^{(1)} \end{pmatrix}, \begin{pmatrix} x \\ a \end{pmatrix} \right\rangle$$

for $y^{(1)}, a^{(1)}, x^{(1)} \in \mathbb{R}$. Eqn. (2) implies

$$\langle y^{(1)}, y_{(1)} \rangle = y^{(1)} \cdot y_{(1)} = \left\langle \begin{pmatrix} a^{(1)} \\ x^{(1)} \end{pmatrix}, \begin{pmatrix} a_{(1)} \\ x_{(1)} \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} a^{(1)} \\ x^{(1)} \end{pmatrix}, \begin{pmatrix} x \\ a \end{pmatrix} \right\rangle \cdot y_{(1)}$$

yielding

$$\begin{pmatrix} a_{(1)} \\ x_{(1)} \end{pmatrix} = \begin{pmatrix} x \\ a \end{pmatrix} \cdot y_{(1)}$$

and hence Eqn. (3).

Lemma 2. *The adjoint of an inner vector product*

$$y = \langle \mathbf{a}, \mathbf{x} \rangle \equiv \mathbf{a}^T \cdot \mathbf{x} = \sum_{i=0}^{n-1} a_i \cdot x_i$$

with active inputs $\mathbf{a} \in \mathbb{R}^n$ *and* $\mathbf{x} \in \mathbb{R}^n$ *yielding the active output* $y \in \mathbb{R}$ *is computed as*

$$\begin{aligned} \mathbf{a}_{(1)} &= \mathbf{x} \cdot y_{(1)} \\ \mathbf{x}_{(1)} &= \mathbf{a} \cdot y_{(1)} \end{aligned} \tag{4}$$

for $y_{(1)} \in \mathbb{R}$ *yielding* $\mathbf{a}_{(1)} \in \mathbb{R}^n$ *and* $\mathbf{x}_{(1)} \in \mathbb{R}^n$.

Proof. Differentiation of $y = \mathbf{a}^T \cdot \mathbf{x}$, for $\mathbf{a} = (a_i)_{i=0,\ldots,n-1}$ and $\mathbf{x} = (x_i)_{i=0,\ldots,n-1}$, with respect to $\mathbf{a}$ and $\mathbf{x}$ yields the tangent

$$\begin{aligned} y^{(1)} &= \sum_{i=0}^{n-1} (x_i \;\; a_i) \cdot \begin{pmatrix} a_i^{(1)} \\ x_i^{(1)} \end{pmatrix} = \sum_{i=0}^{n-1} \left(x_i \cdot a_i^{(1)} + a_i \cdot x_i^{(1)} \right) \\ &= \sum_{i=0}^{n-1} x_i \cdot a_i^{(1)} + \sum_{i=0}^{n-1} x_i^{(1)} \cdot a_i = \mathbf{x}^T \cdot \mathbf{a}^{(1)} + \mathbf{a}^T \cdot \mathbf{x}^{(1)} = (\mathbf{x}^T \;\; \mathbf{a}^T) \cdot \begin{pmatrix} \mathbf{a}^{(1)} \\ \mathbf{x}^{(1)} \end{pmatrix} . \end{aligned}$$

Eqn. (2) implies

$$y_{(1)} \cdot y^{(1)} = (\mathbf{a}_{(1)}^T \;\; \mathbf{x}_{(1)}^T) \cdot \begin{pmatrix} \mathbf{a}^{(1)} \\ \mathbf{x}^{(1)} \end{pmatrix} = y_{(1)} \cdot (\mathbf{x}^T \;\; \mathbf{a}^T) \cdot \begin{pmatrix} \mathbf{a}^{(1)} \\ \mathbf{x}^{(1)} \end{pmatrix}$$

yielding $(\mathbf{a}_{(1)}^T \;\; \mathbf{x}_{(1)}^T) = y_{(1)} \cdot (\mathbf{x}^T \;\; \mathbf{a}^T)$ and hence Eqn. (4).

The following derivation of adjoint matrix-vector and matrix-matrix products relies on serialization of matrices. Individual rows of a matrix $A \in \mathbb{R}^{m \times n}$ are denoted as $\mathbf{a}_i \in \mathbb{R}^{1 \times n}$ for $i = 0, \ldots, m-1$; columns are denoted as $\mathbf{a}^j \in \mathbb{R}^m$

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08754-7_27

for $i = 0, \ldots, n-1$. (Row) Vectors in $\mathbb{R}^{1\times n}$ are denoted as $(v_j)^{j=0,\ldots,n-1}$; (column) vectors in $\mathbb{R}^m$ are denoted as $(v_i)_{i=0,\ldots,m-1}$; Consequently, a row-major serialization of A is given by $\left(\mathbf{a}_i^T\right)_{i=0,\ldots,m-1}$. A column-major serialization of A is given by $\left(\mathbf{a}^j\right)_{j=0,\ldots,n-1}$. Tangents and adjoints of the individual entries of A define

$$A^{(1)} = (\mathbf{a}_i^{(1)})_{i=0,\ldots,m-1} = (a_{i,j}^{(1)})_{i=0,\ldots,m-1}^{j=0,\ldots,n-1}$$

and

$$A_{(1)} = (\mathbf{a}_{(1)i})_{i=0,\ldots,m-1} = (a_{(1)i,j})_{i=0,\ldots,m-1}^{j=0,\ldots,n-1},$$

respectively.

Lemma 3. *The adjoint of a matrix-vector product*

$$\mathbf{y} = A \cdot \mathbf{x} \equiv (\mathbf{a}_i \cdot \mathbf{x})_{i=0,\ldots,m-1}$$

with active inputs $A \in \mathbb{R}^{m\times n}$ and $\mathbf{x} \in \mathbb{R}^n$ yielding the active output $\mathbf{y} \in \mathbb{R}^m$ is computed as

$$\begin{aligned} \mathbf{x}_{(1)} &= A^T \cdot \mathbf{y}_{(1)} \\ A_{(1)} &= \mathbf{y}_{(1)} \cdot \mathbf{x}^T \end{aligned} \tag{5}$$

for $\mathbf{y}_{(1)} \in \mathbb{R}^m$ yielding $\mathbf{x}_{(1)} \in \mathbb{R}^n$ and $A_{(1)} \in \mathbb{R}^{m\times n}$.

Proof. Differentiation of $\mathbf{y} = A \cdot \mathbf{x}$, where $A = (\mathbf{a}_i)_{i=0,\ldots,m-1}$, $\mathbf{x} = (x_j)_{j=0,\ldots,n-1}$ and $\mathbf{y} = (y_i)_{i=0,\ldots,m-1}$, with respect to A and $\mathbf{x}$ yields the tangent

$$\begin{aligned} \mathbf{y}^{(1)} &= \left(\left\langle \begin{pmatrix} \mathbf{x} \\ \mathbf{a}_i^T \end{pmatrix}, \begin{pmatrix} {\mathbf{a}_i^{(1)}}^T \\ \mathbf{x}^{(1)} \end{pmatrix} \right\rangle\right)_{i=0,\ldots,m-1} = \left(\mathbf{x}^T \cdot {\mathbf{a}_i^{(1)}}^T + \mathbf{a}_i \cdot \mathbf{x}^{(1)}\right)_{i=0,\ldots,m-1} \\ &= \left(\mathbf{x}^T \cdot {\mathbf{a}_i^{(1)}}^T\right)_{i=0,\ldots,m-1} + \left(\mathbf{a}_i \cdot \mathbf{x}^{(1)}\right)_{i=0,\ldots,m-1} \\ &= \left(\mathbf{a}_i^{(1)} \cdot \mathbf{x}\right)_{i=0,\ldots,m-1} + \left(\mathbf{a}_i \cdot \mathbf{x}^{(1)}\right)_{i=0,\ldots,m-1} \\ &= \left(\mathbf{a}_i^{(1)}\right)_{i=0,\ldots,m-1} \cdot \mathbf{x} + (\mathbf{a}_i)_{i=0,\ldots,m-1} \cdot \mathbf{x}^{(1)} = A^{(1)} \cdot \mathbf{x} + A \cdot \mathbf{x}^{(1)} . \end{aligned}$$

Eqn. (2) implies

$$\begin{aligned}
\left\langle \mathbf{y}_{(1)}, \mathbf{y}^{(1)} \right\rangle &= \left\langle \begin{pmatrix} \left(\mathbf{a}_{(1)i}^T\right)_{i=0,\ldots,m-1} \\ \mathbf{x}_{(1)} \end{pmatrix}, \begin{pmatrix} \left(\mathbf{a}_i^{(1)T}\right)_{i=0,\ldots,m-1} \\ \mathbf{x}^{(1)} \end{pmatrix} \right\rangle \\
&= \left(\mathbf{a}_{(1)i}^T\right)_{i=0,\ldots,m-1}^T \cdot \left(\mathbf{a}_i^{(1)T}\right)_{i=0,\ldots,m-1} + \mathbf{x}_{(1)}^T \cdot \mathbf{x}^{(1)} \\
&= \mathbf{y}_{(1)}^T \cdot \left(A^{(1)} \cdot \mathbf{x} + A \cdot \mathbf{x}^{(1)}\right) = \mathbf{y}_{(1)}^T \cdot A^{(1)} \cdot \mathbf{x} + \mathbf{y}_{(1)}^T \cdot A \cdot \mathbf{x}^{(1)} \\
&= \left(\left(y_{(1)i} \cdot \mathbf{a}_i^{(1)T}\right)_{i=0,\ldots,m-1}\right)^T \cdot (\mathbf{x})_{i=0,\ldots,m-1} + \mathbf{y}_{(1)}^T \cdot A \cdot \mathbf{x}^{(1)} \\
&= \underbrace{\left(\left(y_{(1)i} \cdot \mathbf{x}\right)_{i=0,\ldots,m-1}\right)^T}_{=\left((\mathbf{a}_{(1)i}^T)_{i=0,\ldots,n-1}\right)^T} \cdot \left(\mathbf{a}_i^{(1)T}\right)_{i=0,\ldots,m-1} + \underbrace{\mathbf{y}_{(1)}^T \cdot A}_{=\mathbf{x}_{(1)}^T} \cdot \mathbf{x}^{(1)} \ ,
\end{aligned}$$

where $(\mathbf{x})_{i=0,\ldots,m-1} \in \mathbb{R}^{m \cdot n}$ denotes a concatenation of m copies of $\mathbf{x} \in \mathbb{R}^n$ as a column vector. Eqn. (5) follows immediately.

Lemma 4. *The adjoint of a matrix-matrix product $Y = A \cdot X$ with active inputs $A \in \mathbb{R}^{m \times p}$, $X \in \mathbb{R}^{p \times n}$ yielding the active output $Y \in \mathbb{R}^{m \times n}$ is computed as*

$$\begin{aligned}
A_{(1)} &= Y_{(1)} \cdot X^T \\
X_{(1)} &= A^T \cdot Y_{(1)}
\end{aligned} \tag{6}$$

for $Y_{(1)} \in \mathbb{R}^{m \times n}$ yielding $A_{(1)} \in \mathbb{R}^{m \times p}$ and $X_{(1)} \in \mathbb{R}^{p \times n}$.

Proof. Differentiation of $Y = A \cdot X$, where $A = (\mathbf{a}_i)_{i=0,\ldots,m-1}$, $X = \left(\mathbf{x}^k\right)^{k=0,\ldots,p-1}$ and $Y = \left(\mathbf{y}^k\right)^{k=0,\ldots,p-1}$, with respect to A and X yields tangents

$$\mathbf{y}^{(1)k} = \left(\left\langle \begin{pmatrix} \mathbf{x}^k \\ \mathbf{a}_i^T \end{pmatrix}, \begin{pmatrix} {\mathbf{a}_i^{(1)}}^T \\ \mathbf{x}^{(1)k} \end{pmatrix} \right\rangle\right)_{i=0,\ldots,m-1} = A^{(1)} \cdot \mathbf{x}^k + A \cdot \mathbf{x}^{(1)k} \ .$$

for $k = 0, \ldots, p-1$ and hence

$$Y^{(1)} = A^{(1)} \cdot X + A \cdot X^{(1)} \ .$$

Eqn. (2) implies

$$\begin{aligned}
\left\langle \mathbf{y}_{(1)}^k, \mathbf{y}^{(1)k} \right\rangle &= \left\langle \begin{pmatrix} \left(\mathbf{a}_{(1)i}^T\right)_{i=0,\ldots,m-1} \\ \mathbf{x}_{(1)k} \end{pmatrix}, \begin{pmatrix} \left(\mathbf{a}_i^{(1)T}\right)_{i=0,\ldots,m-1} \\ \mathbf{x}^{(1)k} \end{pmatrix} \right\rangle \\
&= \underbrace{\left(\left(y_{(1)i}^k \cdot \mathbf{x}^k\right)_{i=0,\ldots,m-1}\right)^T}_{=\left((\mathbf{a}_{(1)i}^T)_{i=0,\ldots,m-1}\right)^T} \cdot \left(\mathbf{a}_i^{(1)T}\right)_{i=0,\ldots,m-1} + \underbrace{\mathbf{y}_{(1)}^{k^T} \cdot A}_{=\mathbf{x}_{(1)}^{k^T}} \cdot \mathbf{x}^{(1)k}
\end{aligned}$$

for $k = 0, \ldots, p-1$ and hence the Eqn. (6).

4 Systems of Linear Equations Revisited

Lemmas 5 and 6 form the basis for the new proof of Eqn. (1).

Lemma 5. *The tangent*

$$Y^{(1)} = A \cdot X^{(1)} \cdot B$$

of $Y = A \cdot X \cdot B$ for active $X \in \mathbb{R}^{n \times q}$, $Y \in \mathbb{R}^{m \times p}$ and passive $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{q \times p}$ implies the adjoint

$$X_{(1)} = A^T \cdot Y_{(1)} \cdot B^T \ .$$

Proof.

$$Y^{(1)} = Z^{(1)} \cdot B \quad \Rightarrow \quad Z_{(1)} = Y_{(1)} \cdot B^T$$

follows from application of Lemma 4 to $Y = Z \cdot B$ with passive B.

$$Z^{(1)} = A \cdot X^{(1)} \quad \Rightarrow \quad X_{(1)} = A^T \cdot Z_{(1)}$$

follows from application of Lemma 4 to $Z = A \cdot X$ with passive A. Substitution of $Z^{(1)}$ and $Z_{(1)}$ yields Lemma 5.

Lemma 6. *The tangent*

$$Y^{(1)} = \sum_{i=0}^{k-1} A_i \cdot X_i^{(1)} \cdot B_i$$

of $Y = \sum_{i=0}^{k-1} A_i \cdot X_i \cdot B_i = \sum_{i=0}^{k-1} Y_i$ with active $X_i \in \mathbb{R}^{n_i \times q_i}$, $Y \in \mathbb{R}^{m \times p}$ and with passive $A_i \in \mathbb{R}^{m \times n_i}$, $B_i \in \mathbb{R}^{q_i \times p}$ implies the adjoint

$$X_{i(1)} = A_i^T \cdot Y_{(1)} \cdot B_i^T$$

for $i = 0, \ldots, k-1$

Proof. From

$$Y_i^{(1)} = A_i \cdot X_i^{(1)} \cdot B_i$$

follows with Lemma 5

$$X_{i(1)} = A_i^T \cdot Y_{i(1)} \cdot B_i^T$$

for $i = 0, \ldots, k-1$. Moreover, $Y^{(1)} = \sum_{i=0}^{k-1} Y_i^{(1)}$ implies $Y_{i(1)} = Y_{(1)}$ due to identity Jacobians of Y with respect to Y_i for $i = 0, \ldots, k-1$ and hence Lemma 6.

Theorem 1. *Adjoints of systems $A \cdot \mathbf{x} = \mathbf{b}$ of n linear equations with invertible $A \in \mathbb{R}^{n \times n}$ and right-hand side $\mathbf{b} \in \mathbb{R}^n$ are evaluated at the primal solution $\mathbf{x} = A^{-1} \cdot \mathbf{b} \in \mathbb{R}^n$ by Eqn. (1).*

DOI: 10.1007/978-3-031-08754-7_27

Proof. Differentiation of $A \cdot \mathbf{x} = \mathbf{b}$ with respect to A and $\mathbf{b}$ yields the tangent system

$$A^{(1)} \cdot \mathbf{x} + A \cdot \mathbf{x}^{(1)} = \mathbf{b}^{(1)}$$

which implies

$$\mathbf{x}^{(1)} = A^{-1} \cdot \mathbf{b}^{(1)} \cdot I_n - A^{-1} \cdot A^{(1)} \cdot \mathbf{x}$$

with identity $I_n \in \mathbb{R}^{n \times n}$. Lemma 6 yields

$$\begin{aligned} \mathbf{b}_{(1)} &= A^{-T} \cdot \mathbf{x}_{(1)} \cdot I_n^T \\ A_{(1)} &= -\underbrace{A^{-T} \cdot \mathbf{x}_{(1)}}_{=\mathbf{b}_{(1)}} \cdot \mathbf{x}^T \end{aligned}$$

and hence Eqn. (1).

5 Conclusion

As observed previously by various authors a possibly available factorization of A can be reused both for the tangent ($A \cdot \mathbf{x}^{(1)} = \mathbf{b}^{(1)} - A^{(1)} \cdot \mathbf{x}$) and the adjoint ($A^T \cdot \mathbf{b}_{(1)} = \mathbf{x}_{(1)}$) systems. The additional worst case computational cost of $O(n^3)$ can thus be reduced to $O(n^2)$. Higher-order tangents [adjoints] of linear systems amount to repeated solutions of linear systems with the same [transposed] system matrix combined with tangent [adjoint] BLAS.

References

1. N. Dunford and J. Schwartz, *Linear Operators. I. General Theory*, With the assistance of W. G. Bade and R. G. Bartle. Pure and Applied Mathematics, Vol. 7, Interscience Publishers, Inc., New York, 1958.
2. M. B. Giles, *Collected matrix derivative results for forward and reverse mode algorithmic differentiation*, in Advances in Automatic Differentiation, C. Bischof, M. Bücker, P. Hovland, U. Naumann, and J. Utke, eds., Springer, 2008, pp. 35–44.
3. A. Griewank and A. Walther, *Evaluating Derivatives. Principles and Techniques of Algorithmic Differentiation, Seocnd Edition*, no. OT105 in Other Titles in Applied Mathematics, SIAM, 2008.
4. C. Lawson, R. Hanson, D. Kincaid, and F. Krogh, *Basic linear algebra subprograms for Fortran usage*, ACM Trans. Math. Softw., 5 (1979), pp. 308–323.
5. U. Naumann, *The Art of Differentiating Computer Programs. An Introduction to Algorithmic Differentiation.*, no. SE24 in Software, Environments, and Tools, SIAM, 2012.