JUIN / JUNE 1999

IN THIS ISSUE / DANS CE NUMÉRO

21
No 2

C. R. Math. Rep. Acad. Sci. Canada Vol. 21 (2), 1999 pp. 33–38

UN ALGORITHME NUMÉRIQUE RAPIDE ET ADAPTATIF À BASE D'ONDELETTES POUR LES ÉQUATIONS D'ÉVOLUTION

GUILLAUME CHIAVASSA ET JACQUES LIANDRAT

Présenté par Vlastimil Dlab, FRSC

ABSTRACT. An adaptive numerical scheme based on wavelets for the resolution of semi-linear evolution equation is presented. Optimality is reached since the complexity and the required storage are proportional to the number of degrees of freedom. Numerical tests are provided for the Burgers equation with small viscosity ($\nu = 10^{-7}$). The corresponding results are accurate and unreachable with non fully adaptive methods.

RÉSUMÉ. Nous présentons un algorithme adaptatif à base d'ondelettes pour la résolution d'équations d'évolution semi-lineaires. Cet algorithme est optimal puisque sa complexité et la taille mémoire qui est associée à son implémentation, sont proportionnelles au nombre de degrés de liberté du problème. Des tests numériques sont développés en dimension un sur une équation de Burgers à très faible coefficient régularisant ($\nu = 10^{-7}$). On obtient des résultats précis et rapides, déjà hors de portée des méthodes non adaptatives ou non totalement adaptatives, et difficilement atteignables par les méthodes adaptatives classiques.

1. Introduction. Cette note s'inscrit dans un travail de définition et d'implémentation d'algorithmes précis et rapides à base d'ondelettes pour la résolution numérique adaptative d'équations non-linéaires. Nous y montrons que l'utilisation de certaines bases d'ondelettes pour la représentation adaptée de fonctions présentant localement des forts gradients, combinée avec des approximations performantes des opérateurs non-linéaires et d'évolution linéaire conduit à des algorithmes numériques effectivement adaptatifs, c'est-à-dire ayant une complexité et demandant une place mémoire toutes deux proportionnelles au nombre de degrés de liberté du problème. Nous présentons des résultats de tests numériques sur l'équation de Burgers à très faible coefficient régularisant ($\nu = 10^{-7}$).

2. Discrétisation de l'équation et algorithme. Nous nous intéressons à la résolution numérique de l'équation suivante:

$$\text{(1)} \qquad \begin{cases} \partial_t U + G(U) = \nu\, \partial_{xx} U \\ (x,t) \in]0,1[\times [0,T] \\ U(0,t) = U(1,t) \\ U(x,0) = g(x) \end{cases}$$

Reçu par les éditeurs le 1er août, 1998.

où $G(U)$ est un opérateur non-linéaire de la forme $U \mapsto G \circ U$ et ν une constante.

L'approche que nous utilisons pour la discrétisation de l'équation n'est pas standard. Elle consiste tout d'abord à discrétiser l'équation par rapport au temps puis à approcher ensuite le problème elliptique posé à chaque pas de temps en utilisant les bases d'ondelettes [3].

2.1. *Discrétisation temporelle.* Ayant introduit une segmentation $(t_n)_{0 \le n \le N}$ de l'intervalle $[0, T]$, nous savons que U vérifie:

$$U(x, t_{n+1}) = e^{\nu \delta t_n \partial_{xx}} U(x, t_n) - \int_0^{\delta t_n} e^{\nu(\delta t_n - \alpha)\partial_{xx}} G\big(U(t_n + \alpha)\big)\, d\alpha \tag{2}$$

avec $\delta t_n = t_{n+1} - t_n$. Le schéma numérique que nous utilisons s'écrit :

$$U^{n+1}(x) = e^{\nu \delta t_n \partial_{xx}} U^n(x) - \delta t_n \sum_{i=0}^{r} q_{i,n} e^{\nu(\delta t_n + t_n - t_{n-i})\partial_{xx}} G\big(U^{n-i}(x)\big) \tag{3}$$

où $\{q_{i,n}, 0 \le i \le r\}$ sont les coefficients de la formule de quadrature utilisée pour approcher l'intégrale dans (2). Pour simplifier nous écrirons : $U^{n+1}(x) = \sum_{i=0}^{r} L(\delta t_n, t_n, t_{n-i}) U^{n-i}(x)$.

2.2. *Approximation spatiale.* Nous obtenons une approximation de U^{n+1} par un algorithme analysé dans [3] et s'écrivant :

$$\tilde{U}^{n+1} = \sum_{i=0}^{r} \Pi_{V^{(n+1)}} L(\delta t_n, t_n, t_{n-i}) \tilde{U}^{n-i}.$$

L'espace d'approximation au temps t_n, $V^{(n)}$, est choisi de la forme :

$$V^n = V_{j_0} \oplus \sum_{j=j_0}^{p-1} X_j \tag{4}$$

avec $V_{j_0} = \mathrm{vect}\{\phi_{j_0,k}, 0 \le k \le j^{j_0} - 1\}$ et $X_j = \mathrm{vect}\{\psi_{j,k}, k_1(j) \le k \le k_2(j)\}$, et où $\phi_{j_0,k}$ (resp. $\psi_{j,k}$) sont les fonctions d'echelles (resp. les ondelettes) d'une analyse multirésolution périodique de $L^2([0,1])$ [5]. Les indices j_0 et p ainsi que les fonctions k_1 et k_2 dépendent de n. L'approximation $\tilde{U}^{n+1}$ s'écrit alors

$$\begin{aligned} \tilde{U}^{n+1} = \sum_{i=0}^{r} \Big(& \Pi_{V^{(n+1)}} L(\delta t_n, t_n, t_{n-i}) \Pi_{V_{j_0}} \tilde{U}^{n-i} \\ & + \sum_{j=j_0}^{p-1} \Pi_{V^{(n+1)}} L(\delta t_n, t_n, t_{n-i}) \Pi_{X_j} \tilde{U}^{n-i} \Big). \end{aligned} \tag{5}$$

Nous montrons dans la suite qu'un algorithme de complexité $O(\#V^{(n)})$ et ne manipulant que des vecteurs de taille proportionnelle à $\#V^{(n)}$ ou $\#V^{(n+1)}$ est associé à (5). Les difficultés étant essentiellement liées au calcul du terme non-linéaire ($U \mapsto G(U)$) et à celui de l'action de l'opérateur de la chaleur ($U \mapsto e^{\nu \delta t_n \partial_{xx}} U$), nous nous concentrons maintenant sur ces deux parties.

2.3. *Calcul du terme non-linéaire.* Pour $PG_q(U) = 2^{-q/2}\sum_k G(2^{q/2}\langle U, \phi_{q,k}\rangle)\phi_{q,k}$, si l'on pose

$$PG_{V^{(n)}}(U) = PG_{j_0}(U) + \sum_{j=j_0}^{p-1} \Pi_{X_j} PG_{j+1}(U) \tag{6}$$

nous savons, ([4]), que $\left\|\Pi_{V^{(n)}}\left(G(U)\right) - PG_{V^{(n)}}(U)\right\|_{L^2} = O(2^{-(n_\phi+1)j_0})$ où n_ϕ est le nombre de moments nuls[1] de la fonction ϕ. Ainsi, pour n_ϕ et j_0 assez grands, l'estimation PG peut remplacer G. Connaissant la décomposition de U^n dans $V^{(n)}$, le calcul de $PG_{V^{(n)}}$ s'effectue alors en utilisant les algorithmes pyramidaux classiques de la transformée en ondelettes. Lorsque l'espace $V^{(n)}$ est c-structuré, *i.e.*, vérifie

$$\begin{aligned} &\exists c \text{ tel que } \forall j, \quad j_0 < j \le p-1, \quad \psi_{j,k} \in X_j \Rightarrow \psi_{j-1,k'} \in X_{j-1} \\ &\qquad\qquad \text{si} \quad |2k' - k| \le c \quad c \ge lg \end{aligned} \tag{7}$$

la complexité de ces algorithmes est $O(\#V^{(n)})$ (lg représente la longueur des filtres de l'analyse multirésolution).

2.4. *Application de l'opérateur de la chaleur* L'algorithme que nous utilisons pour le calcul de $\Pi_{V^{(n)}}C(U)$ où $C(U) = e^{\nu\delta t_n \partial_{xx}}U$ est le suivant :

$$\begin{aligned} \text{Si} \quad U &= \sum_{k=0}^{2^{j_0}-1} c_{j_0,k}\phi_{j_0,k} + \sum_{j\ge j_0,k} d_{j,k}\psi_{j,k}, \\ \text{alors} \quad C(U) &= \sum_{k=0}^{2^{j_0}-1} c_{j_0,k}\tau_{j_0,k} + \sum_{j\ge j_0,k} d_{j,k}\theta_{j,k}, \end{aligned} \tag{8}$$

avec $\tau_{j_0,k} = e^{\nu\delta t_n\partial_{xx}}\phi_{j_0,k}$ et $\theta_{j,k} = e^{\nu\delta t_n\partial_{xx}}\psi_{j,k}$.

Suivant [5], les fonctions $\theta_{j,k}$ sont des "vaguelettes" c'est-à-dire des fonctions localisées uniformément par rapport à j de même que leur transformée de Fourier. Il s'ensuit que la projection de ces fonctions sur la base d'ondelettes peut être calculée de façon pyramidale puisque les produits scalaires $\langle\theta_{j,k}, \psi_{j',k'}\rangle$ vérifient pour ϵ' donné :

$$\begin{aligned} &\exists (l,d) \in \mathbb{N} \text{ tels que } \quad |\langle\theta_{j,k}, \psi_{j',k'}\rangle| \le \epsilon' \quad \text{pour} \quad |j - j'| > l \\ &\qquad\qquad\qquad\qquad\qquad\quad |\langle\theta_{j,k}, \psi_{j,k'}\rangle| \le \epsilon' \quad \text{pour} \quad |k - k'| > d. \end{aligned} \tag{9}$$

Du fait de la décroissance des fonctions mises en jeu, la complexité numérique du calcul des produits scalaires (9), qui augmente théoriquement avec j en $O(j2^j)$, peut être rendue constante à partir d'un certain seuil j_c dépendant seulement de l'ondelette ψ [3] ($j_c = 9$ dans la partie (3)).

L'implémentation de la projection de $C(U)$ sur $V^{(n)}$ conduit donc réellement à un algorithme de complexité $O(\#V^{(n)})$ et le calcul des produits scalaires (9) est de complexité $O(p)$, p étant défini en (4).

[1] Classiquement, nous appelons moment d'ordre α ($\alpha \in \mathbb{N}$) de la fonction f la quantité $\int x^\alpha f(x)\,dx$ quand elle existe.

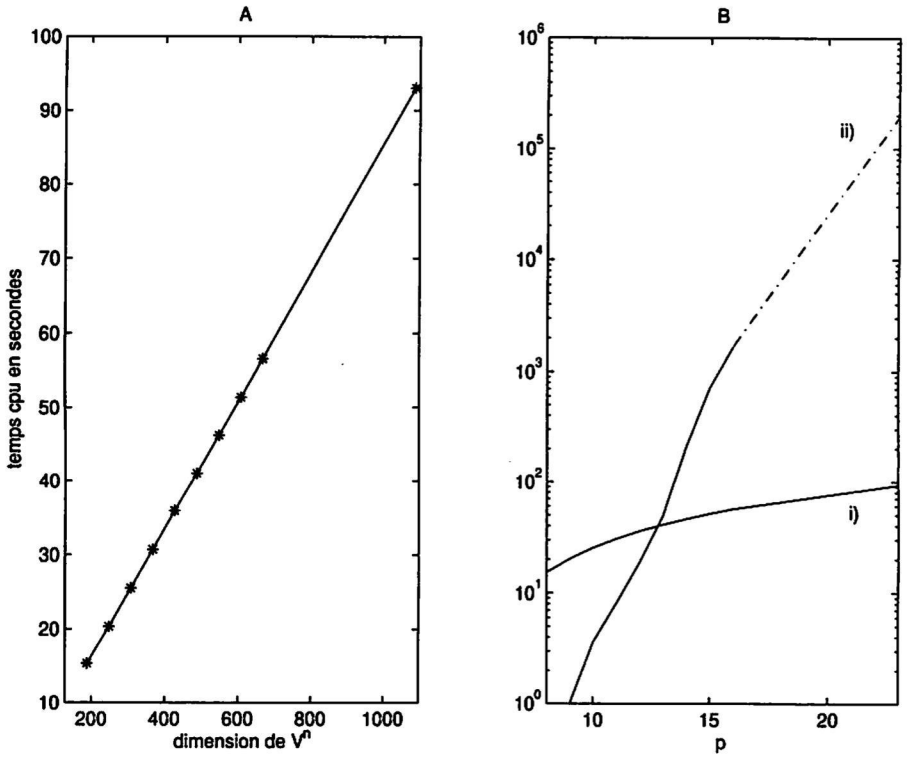

Figure 1: A) Temps de calcul pour 1000 itérations de l'algorithme complet en fonction de la dimension de l'espace d'approximation $V^{(n)}$. B) i) Temps de calcul pour le même test que sur la figure A) en fonction de p. ii) Temps de calcul obtenus (trait plein) et estimés (pointillés) pour une méthode spectrale non-adaptative.

3. Résultats numériques. Ils concernent l'équation de Burgers avec un très faible coefficient régularisant ν :

$$\begin{cases} \partial_t U + \frac{1}{2}\partial_x(U^2) = \nu \partial_{xx} U \\ (x,t) \in]0,1[\times [0,T] \\ U(0,t) = U(1,t) \\ U(x,0) = \sin(2\pi x). \end{cases} \tag{10}$$

De nombreux algorithmes, ([1], [2] par exemple), ont déjà été testés sur cette équation avec des valeurs de la viscosité ν de l'ordre de 10^{-3} ce qui correspond à une valeur maximale du gradient $\|\partial_x U(x,1/4)\|_\infty \simeq 500$. Nous présentons ici les résultats obtenus pour $\nu = 10^{-7}$ conduisant à $\|\partial_x U(x,1/4)\|_\infty \simeq 5 \times 10^6$. Il s'agit donc déjà d'un problème difficile à approcher (même si la dimension est 1) où l'intérêt de méthodes rapides et adaptatives peut être mise en évidence. L'algorithme précédent a été implémenté avec un schéma de discrétisation temporel d'ordre 1 à pas de temps variable et des ondelettes splines d'ordre 6. Tous les

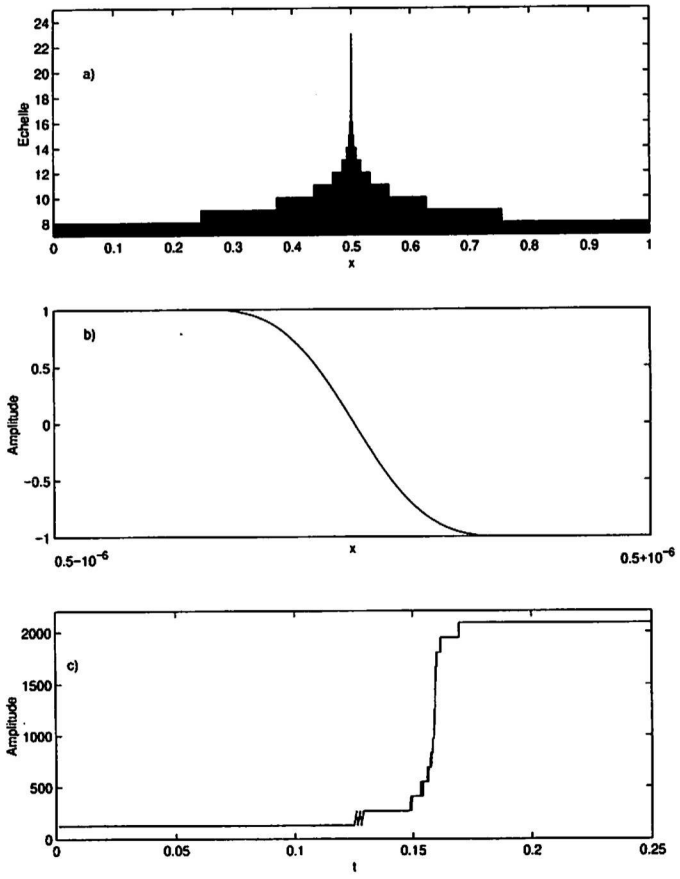

Figure 2: a) $V^{(n_{\max})}$; si $\psi_{j,k} \in V^{(n_{\max})}$, le rectangle $\left[(k2^{-j}, j), (k2^{-j}, j+1), \right.$ $\left.((k+1)2^{-j}, j+1), ((k+1)2^{-j}, j)\right]$ est coloré en noir. b) Zoom de $\tilde{U}^n(x, t_{n_{\max}})$ pour $|x - 0.5| \leq 10^{-6}$. c) Evolution de $V^{(n)}$ en fonction de t.

filtres ont été tronqués à la précision de 10^{-6} ($lg = 108$). La figure 1 A) représente le temps de calcul mesuré sur un ordinateur de type PC200Mhz en fonction de la dimension de $V^{(n)}$ et pour un nombre fixé d'itérations de l'algorithme complet. Notons que ces résultats prennent en compte le temps de calcul des produits scalaires (9) et démontrent ainsi l'optimalité de l'algorithme implémenté.

Sur la figure 1 B), nous comparons ces temps de calcul, (a)), avec ceux donnés par un algorithme équivalent [2] sur une grille régulière de dimension 2^p (b)).

La figure 2 a) représente l'espace d'approximation $V^{(n_{\max})}$ utilisé pour calculer la solution de pente maximale au temps $t_{n_{\max}} = 0.2498$ (on a alors $j_0 = 8, p = 23, k_2 - k_1 = 126$). La valeur estimée de $\|\partial_x \tilde{U}\|_\infty$ est alors 4 514 723, soit une

[2] Méthode pseudo-spectrale classique à base d'exponentielles.

erreur relative de 9×10^{-2} par rapport à la valeur asymptotique ($4 - \frac{1}{2\nu}$ au temps $t = 0.25$). Notons qu'aucune oscillation ne se produit dans la zone de fort gradient (figure 2 b)). L'adaptation de l'espace d'approximation $V^{(n)}$ est automatique et basée sur la norme des coefficients d'ondelettes d'indice $p - 1$. Nous avons représenté sur la figure 2 c) l'évolution de la dimension de $V^{(n)}$ en fonction du temps. La valeur maximale est de 2146, valeur qui est à comparer à 2^{23}, le nombre minimal de degrés de liberté à prendre en compte si aucune technique d'adaptativité n'est utilisée.

Conclusion. L'algorithme à base d'ondelettes que nous avons présenté est totalement adaptatif en terme de nombre d'opérations mais aussi de taille mémoire. Il permet d'obtenir des résultats très précis pour des configurations particulièrement difficiles (solutions à évolution temporelle rapide et présentant localement des gradients spatiaux très forts). De telles configurations sont très difficilement simulables (voire non simulables) par des méthodes classiques et représentent donc un domaine où les méthodes réellement adaptatives à base d'ondelettes sont numériquement intéressantes.

References

[1] C. Basdevant, M. Deville, P. Haldenwang, J. M. Lacroix, J. Ouazzani and R. Peyret, *Spectral and finite difference solution of the Burgers equation.* Computers and Fluids (1) **14**(1986), 23–41.

[2] G. Beylkin and J. Keiser, *On the adaptive numerical solution of nonlinear partial differential equations in wavelet bases.* J. Comput. Phys. (2) **132**(1997), 233–259.

[3] G. Chiavassa, *Algorithmes Adaptatifs en Ondelettes pour la Résolution d'Équations aux Dérivées Partielles.* Thèse, Université d'Aix-Marseille II, Institut de Recherche sur les Phénomènes Hors Equilibre, Juin, 1997.

[4] J. Liandrat and Ph. Tchamitchian, *On the fast approximation of some non linear operators in adapted wavelet spaces.* Adv. Comput. Math. **8**(1998), 179–192.

[5] Y. Meyer, *Ondelettes et Opérateurs : vol. I et II.* Hermann, Paris, 1990.

IRPHE, UMR 6594, La Jetée-Technopôle de Château-Gombert
38, rue F. Joliot-Curie
13451 Marseille Cedex 20
France
email: cassa@marius.univ-mrs.fr

C. R. Math. Rep. Acad. Sci. Canada Vol. 21 (2), 1999 pp. 39–45

ON CONDITIONS FOR $R[\alpha, \alpha^{-1}]$ TO BE EXCLUSIVE AND CONTRACTIONS OF PRINCIPAL IDEALS OF $R[\alpha] \cup R[\alpha^{-1}]$

SUSUMU ODA AND KEN-ICHI YOSHIDA

Presented by Vlastimil Dlab, FRSC

ABSTRACT. Let R be a Noetherian domain with quotient field K and let α denote a non-zero algebraic element over K. We examine what conditions are required for $R[\alpha, \alpha^{-1}] \cap K = R$, determine a contraction $\beta(R[\alpha] \cap R[\alpha^{-1}]) \cap R$ for $\beta \in R[\alpha] \cap R[\alpha^{-1}]$, and investigate when a sequence $\beta, \gamma \in R[\alpha] \cap R[\alpha^{-1}]$ is a regular sequence on $R[\alpha] \cap R[\alpha^{-1}]$ by means of a contraction $(\beta, \gamma)(R[\alpha] \cap R[\alpha^{-1}]) \cap R$.

Let R be a Noetherian domain and $R[X]$ a polynomial ring. Let α be an element of an algebraic field extension L of the quotient field K of R and let $\pi: R[X] \to R[\alpha]$ be the R-algebra homomorphism sending X to α. Let $\varphi_\alpha(X)$ be the monic minimal polynomial of α over K with $\deg \varphi_\alpha(X) = d$ and write $\varphi_\alpha(X) = X^d + \eta_1 X^{d-1} + \cdots + \eta_d$. Then η_i $(1 \leq i \leq d)$ are uniquely determined by α.

Let $I_{\eta_i} := R :_R \eta_i$ and $I_{[\alpha]} := \bigcap_{i=1}^{d} I_{\eta_i}$, the latter of which is called a *generalized denominator ideal* of α. It is easy to see that $I_{[\alpha]} = R[X] :_R \varphi_\alpha(X)$. We say that α is an *anti-integral element* over R if $\operatorname{Ker} \pi = I_{[\alpha]} \varphi_\alpha(X) R[X]$ (*cf.* [OY1], [OSaY]). For $f(X) \in R[X]$, let $C(f(X))$ denote the ideal of R generated by the coefficients of $f(X)$. For an ideal J of $R[X]$, let $C(J)$ denote the ideal generated by the coefficients of the elements in J. If α is an anti-integral element, then $C(\operatorname{Ker} \pi) = C(I_{[\alpha]} \varphi_\alpha(X) R[X]) = I_{[\alpha]}(1, \eta_1, \ldots, \eta_d)$. Put $J_{[\alpha]} := I_{[\alpha]}(1, \eta_1, \ldots, \eta_d)$. Let $\tilde{I}_{[\alpha]} := \bigcap_{i=1}^{d-1} I_{\eta_i}$ and let $\tilde{J}_{[\alpha]} := I_{[\alpha]}(1, \eta_1, \ldots, \eta_{d-1})$.

If $J_{[\alpha]} \not\subseteq p$ for all $p \in \operatorname{Dp}_1(R) := \{p \in \operatorname{Spec}(R) \mid \operatorname{depth} R_p = 1\}$, then α is called a *super-primitive* element over R.

It is known that a super-primitive element is an anti-integral element [OSaY, (1.12)].

It is also known that any algebraic element over a Krull domain R is anti-integral over R [OSaY, (1.13)]. Put $B_\alpha := R[\alpha] \cap R[\alpha^{-1}]$.

When α is an element in K, $\varphi_\alpha(X) = X - \alpha$. So we have $J_{[\alpha]} = I_{[\alpha]}(1, \alpha) = I_\alpha(1, \alpha) = I_\alpha + \alpha I_\alpha = I_\alpha + I_{\alpha^{-1}}$, where $I_\alpha := R :_R \alpha$, a *denominator ideal* of $\alpha \in K$.

Our objective of this paper is to investigate the following problems:

Received by the editors 8 December, 1998.
AMS subject classification: Primary: 13C20; secondary: 13F99.

(i) What conditions are required for $R[\alpha, \alpha^{-1}] \cap K = R$;
(ii) Determine $\beta B_\alpha \cap R$ for $\beta \in B_\alpha$;
(iii) Let $\beta, \gamma \in B_\alpha$, and examine some conditions for β, γ to be a regular sequence on B_α.

In this paper, we fix the notations and the definitions mentioned above unless otherwise specified.

Our general reference for unexplained technical terms is [M].

1. Conditions for the extensions $R[\alpha, \alpha^{-1}]$ to be exclusive. Note that $\varphi_\alpha(X) = X^d + \eta_1 X^{d-1} + \cdots + \eta_d$ with $\eta_j \in K$, where $\eta_0 := 1$, and put $\varphi_{\alpha^{-1}}(X) = X^d + \eta'_1 X^{d-1} + \cdots + \eta'_d$ with $\eta'_j \in K$, where $\eta'_0 := 1$. Then $\eta'_j = \eta_{d-j}/\eta_d$ for every j $(0 \le j \le d)$.

Put $\tilde{J}^{(i)}_{[\alpha]} := I_{[\alpha]}(\eta_0, \eta_1, \ldots, \eta_{i-1}, \eta_{i+1}, \ldots, \eta_d)$ for $(0 \le i \le d)$, and put $\tilde{J}^{(i)}_{[\alpha^{-1}]} := I_{[\alpha^{-1}]}(\eta'_0, \eta'_1, \ldots, \eta'_{i-1}, \eta'_{i+1}, \ldots, \eta'_d)$ for $(0 \le i \le d)$. It is clear that $\tilde{J}^{(d)}_{[\alpha]} = \tilde{J}_{[\alpha]}$ and $\tilde{J}^{(d)}_{[\alpha^{-1}]} = \tilde{J}_{[\alpha^{-1}]}$ by definition.

We start with the following lemma.

LEMMA 1.1. *The following equalities hold:*
(i) $I_{\eta'_i} = \eta_d I_{\eta_{d-i}}$ *for all* i $(0 \le i \le d)$;
(ii) $I_{[\alpha^{-1}]} = \eta_d I_{[\alpha]}$;
(iii) $J_{[\alpha]} = J_{[\alpha^{-1}]}$;
(iv) $\tilde{J}^{(i)}_{[\alpha]} = \tilde{J}^{(d-i)}_{[\alpha^{-1}]}$ *for each* i $(0 \le i \le d)$.

PROOF. (i) Take $x \in \eta_d I_{\eta_{d-i}}$. Then $x = \eta_d y$ with $y \in I_{\eta_{d-i}}$ and $x\eta_{d-i} = \eta_{d-i} y \eta_d \in \eta_d R$. Hence $x(\eta_{d-i}/\eta_d) \in R$, which implies that $x \in I_{\eta_{d-i}/\eta_d} = I_{\eta'_i}$. Conversely take $x \in I_{\eta'_{d-i}}$. Hence $x\eta'_{d-i} = x(\eta_i/\eta_d) = (x/\eta_d)\eta_i \in R$ and so $x/\eta_d \in I_{\eta_i}$. Thus $x \in \eta_d I_{\eta_i}$.

(ii) $I_{[\alpha^{-1}]} = \bigcap_{j=0}^{d} I_{\eta'_j} = \bigcap_{j=0}^{d} \eta_d I_{\eta_j} = \eta_d I_{[\alpha]}$.

(iii) $J_{[\alpha]} = I_{[\alpha]}(\eta_0, \ldots, \eta_d) = \eta_d I_{[\alpha]}(\eta_0/\eta_d, \ldots, \eta_{d-1}/\eta_d, \eta_d/\eta_d) = I_{[\alpha^{-1}]}(\eta'_0, \ldots, \eta'_d) = J_{[\alpha^{-1}]}$.

(iv) is obtained by the same way as the proof of (ii). ∎

LEMMA 1.2 (*cf.* [OY2, THEOREM 5 AND ITS PROOF]). *If α is exclusive over R, i.e., $R[\alpha] \cap K = R$, then either* $\operatorname{grade}(\tilde{J}_{[\alpha]}) > 1$ *or* $\tilde{J}_{[\alpha]} = R$.

Recall that if α is a super-primitive element (resp. an anti-integral element) over R then so is α^{-1} over R (*cf.* [KY] and Lemma 1.1).

PROPOSITION 1.3.. *Assume that α is exclusive over R, i.e., $R[\alpha] \cap K = R$ and that R contains an infinite field. Then α (resp. α^{-1}) is exclusive over R, i.e., $R[\alpha] \cap K = R$ (resp. $R[\alpha^{-1}] \cap K = R$) if and only if either* $\operatorname{grade}(\tilde{J}_{[\alpha]}) > 1$ *(resp.* $\operatorname{grade}(\tilde{J}_{[\alpha^{-1}]}) > 1$*) or* $\tilde{J}_{[\alpha]} = R$ *(resp.* $\tilde{J}_{[\alpha^{-1}]} = R$*).*

PROOF. The implication ($\Rightarrow$) follows from Lemma 1.2.

The reverse implication ($\Leftarrow$) follows from [OY2, Theorem 5]. ∎

PROPOSITION 1.4. *Assume that α is a super-primitive element over R. Then the following statements are equivalent for each i $(0 \leq i \leq d)$:*

(i) $\bigcap_{j \neq i} I_{\eta_j} \subseteq I_{\eta_i}$;

(i') $\bigcap_{j \neq d-i} I_{\eta'_j} \subseteq I_{\eta'_{d-i}}$;

(ii) $\operatorname{grade}(\tilde{J}^{(i)}_{[\alpha]}) > 1$ *or* $\tilde{J}^{(i)}_{[\alpha]} = R$;

(ii') $\operatorname{grade}(\tilde{J}^{(d-i)}_{[\alpha^{-1}]}) > 1$ *or* $\tilde{J}^{(d-i)}_{[\alpha^{-1}]} = R$.

PROOF. The equivalences: (i) $\Leftrightarrow$ (i') and (ii) $\Leftrightarrow$ (ii') are shown in Lemma 1.1.

(i) $\Rightarrow$ (ii): We assume that $\tilde{J}^{(i)}_{[\alpha]} \neq R$. Suppose that $\operatorname{grade}(\tilde{J}^{(i)}_{[\alpha]}) = 1$. Then there exists $p \in \operatorname{Dp}_1(R)$ such that $\tilde{J}^{(i)}_{[\alpha]} = I_{[\alpha]}(\eta_1, \ldots, \eta_{i-1}, \eta_{i+1}, \ldots, \eta_d) \subseteq p$. Since $p \in \operatorname{Dp}_1(R)$, there exists an element $\beta \in K$ such that $I_\beta = p$. Thus $\beta J^{(i)}_{[\alpha]} = \beta I_{[\alpha]}(\eta_1, \ldots, \eta_{i-1}, \eta_{i+1}, \ldots, \eta_d) \subseteq R$. Hence $(I_{[\alpha]}\beta)\eta_j \subseteq R$ for all $j \neq i$, which yields that $I_{[\alpha]}\beta \subseteq \bigcap_{j \neq i} I_{\eta_j}$. Since $\bigcap_{j \neq i} I_{\eta_j} \subseteq I_{\eta_i}$, we have that $I_{[\alpha]}\beta\eta_i \subseteq R$. Thus $J_{[\alpha]}\beta \subseteq R$, that is, $J_{[\alpha]} \subseteq I_\beta = P$. But since α is super-primitive over R, we have $J_{[\alpha]} \not\subseteq p$, which is a contradiction. So we conclude that $\operatorname{grade}(\tilde{J}^{(i)}_{[\alpha]} > 1$.

(ii) $\Rightarrow$ (i): First consider the case $\tilde{J}^{(i)}_{[\alpha]} \neq R$. Suppose that there exists an element $a \in \bigcap_{j \neq i} I_{\eta_j}$ but $a \notin I_{\eta_i}$. Then $\beta := a\eta_i \notin R$. So $\beta I_{[\alpha]}(\eta_0, \eta_1, \ldots, \eta_{i-1}, \eta_{i+1}, \ldots, \eta_d) \subseteq R$. Since $\operatorname{grade}(\tilde{J}^{(i)}_{[\alpha]}) = \operatorname{grade}(I_{[\alpha]}(\eta_0, \eta_1, \ldots, \eta_{i-1}, \eta_{i+1}, \ldots, \eta_d)) > 1$, we have $\beta R_p = \beta(I_{[\alpha]}(\eta_0, \eta_1, \ldots, \eta_{i-1}, \eta_{i+1}, \ldots, \eta_d)R_p \subseteq R_p$ for every $p \in \operatorname{Dp}_1(R)$. Hence $\beta \in \bigcap_{p \in \operatorname{Dp}_1(R)} R_p = R$, which is a contradiction. Thus $\bigcap_{j \neq i} I_{\eta_j} \subseteq I_{\eta_i}$. Second, suppose that $\tilde{J}^{(i)}_{[\alpha]} = R$. Then there exists an equality: $1 = a_0\eta_0 + a_1\eta_1 + \cdots + a_{i-1} + a_{i+1}\eta_{i+1} + \cdots + a_d\eta_d$ with $a_i \in I_{[\alpha]}$. So $\eta_i = (a_0\eta_i)\eta_0 + (a_1\eta_i)\eta_1 + \cdots + (a_{i-1}\eta_i)\eta_{i-1} + (a_{i+1}\eta_i)\eta_{i+1} + \cdots + (a_d\eta_i)\eta_d$. Take $X \in \bigcap_{j \neq i} I_{\eta_j}$. Then $x\eta_i = x(a_0\eta_i)\eta_0 + x(a_1\eta_i)\eta_1 + \cdots + x(a_{i-1}\eta_i)\eta_{i-1} + x(a_{i+1}\eta_i)\eta_{i+1} + \cdots + x(a_d\eta_i)\eta_d \in R$, and hence $x \in I_{\eta_i}$. Therefore we conclude that $\bigcap_{j \neq i} I_{\eta_j} \subseteq I_{\eta_i}$.

The equivalence (i') $\Leftrightarrow$ (ii') can be shown by the same way as the proof of (i) $\Leftrightarrow$ (ii). ∎

REMARK 1.5. Let A be an integral domain containing R. Then the following statements are equivalent:

(1) A is exclusive over R (*i.e.*, $A \cap K = R$);

(2) A_p is exclusive over R_p (*i.e.*, $A_p \cap K = R_p$) for all $p \in \operatorname{Dp}_1(R)$.

Indeed, we have only to recall that $\bigcap_{p \in \operatorname{Dp}_1(R)} R_p = R$.

THEOREM 1.6. *Assume that α is a super-primitive element of degree d over R, that R contains an infinite field and that* $\operatorname{grade}(I_{[\alpha]} + I_{[\alpha^{-1}]}) > 1$. *Then the following statements are equivalent:*

(i) $R[\alpha, \alpha^{-1}]$ *is exclusive over R, i.e., $R[\alpha, \alpha^{-1}] \cap K = R$;*

(ii) $\operatorname{grade}(\tilde{J}^{(i)}_{[\alpha]}) > 1$ *or* $\tilde{J}^{(i)}_{[\alpha]} = R$ *for every i $(0 \leq i \leq d)$;*

(iii) $\bigcap_{j \neq i} I_{\eta_j} \subseteq I_{\eta_i}$ *for every i $(0 \leq i \leq d)$;*

(ii') $\operatorname{grade}(\tilde{J}^{(i)}_{[\alpha^{-1}]}) > 1$ *or* $\tilde{J}^{(i)}_{[\alpha^{-1}]} = R$ *for every* i $(0 \le i \le d)$;
(iii') $\bigcap_{j\neq i} I_{\eta'_j} \subseteq I_{\eta'_i}$ *for every* i $(0 \le i \le d)$;

PROOF. The equivalences (ii) $\Leftrightarrow$ (iii), (ii') $\Leftrightarrow$ (iii') and (ii) $\Leftrightarrow$ (ii') follow from Proposition 1.4.

(i) $\Rightarrow$ (iii): Suppose that $\bigcap_{j\neq i} I_{\eta_j} \not\subseteq I_{\eta_i}$ for some i $(0 \le i \le d)$. Then there exists an element $a \in \bigcap_{j\neq i} I_{\eta_j}$ with $\lambda := a\eta_i \notin R$. Since $\varphi_\alpha(\alpha) = 0$, we have $a\varphi_\alpha(\alpha) = a\alpha^d + a\eta_1\alpha^{d-1} + \cdots + a\eta_i\alpha^{d-i} + \cdots + a\eta_d = 0$, which yields $-\lambda\alpha^{d-i} = a\eta_0\alpha^d + \cdots + a\eta_{i-1}\alpha^{d-i+1} + a\eta_{i+1}\alpha^{d-i-1} + \cdots + a\eta_d$. So we have $-\lambda = a\alpha^i + \cdots + (a\eta_{i-1})\alpha + (a\eta i + 1)\alpha^{-1} + \cdots + (a\eta_d)\alpha^{i-d} \in R[\alpha, \alpha^{-1}] \cap K = R$, which is a contradiction. Therefore we conclude that $\bigcap_{j\neq i} I_{\eta_j} \subseteq I_{\eta_i}$ for every i $(0 \le i \le d)$.

(ii') $\Rightarrow$ (i): Since $R[\alpha, \alpha^{-1}] \cap K = R \Leftrightarrow R_p[\alpha, \alpha^{-1}] \cap K = R_p$ for all $p \in \mathrm{Dp}_1(R)$ by Remark 1.5, we may assume that R is a local domain with maximal ideal p with $\operatorname{depth}(R) = 1$. In this case, $I_{[\alpha]} + I_{[\alpha^{-1}]} = R$ because $\operatorname{grade}(I_{[\alpha]} + I_{[\alpha^{-1}]}) > 1$, and hence $I_{[\alpha]} = R$ or $I_{[\alpha^{-1}]} = R$.

Take $\lambda \in R[\alpha, \alpha^{-1}] \cap K$. Then

$$(*) \qquad \lambda = a_0\alpha^n + \cdots + a_{n-1}\alpha + a_n + b_1\alpha^{-1} + \cdots + b_m\alpha^{-m},$$

where $a_i, b_j \in R$. Since $\lambda \in R \Leftrightarrow \lambda - a_n \in R$, we may assume that $a_n = 0$. Put $\psi(X) := a_0X^{n+m} + \cdots + a_{n-1}X^{m+1} - \lambda X^m + b_1X^{m-1} + \cdots + b_m$. Then $\psi(\alpha) = 0$.

Since α^{-1} is a super-primitive element by Lemma 1.1, we may assume that $I_{[\alpha]} = R$ by symmetry. Thus $\varphi_\alpha(X) \in R[X]$.

(a) Assume that $n = 0$. Then $\lambda \in R[\alpha^{-1}]$. Since $\tilde{J}_{[\alpha^{-1}]} = \tilde{J}^{(d)}_{[\alpha^{-1}]} = R$, it follows that α^{-1} is exclusive over R by Proposition 1.3. Hence $\lambda \in R[\alpha^{-1}] \cap K = R$.

(b) Assume that $m \ge d$. Then $\deg \psi(X) = n + m \ge d$. Since $\varphi_\alpha(X) \in R[X]$, there exists a polynomial $h(X) \in R[X]$ such that $\psi^*(X) := \psi(X) - \varphi_\alpha(X)h(X)$ is of degree m. Note that $\psi(X) \in R[X]$ if and only if $\psi^*(X) \in R[X]$. Hence using $\psi^*(X)$ instead of $\psi(X)$, we can assume that $n = 0$, which case is done in (a).

(c) Assume that $m < d$. Then since $\varphi_\alpha(X) \in R[X]$, there exists a polynomial $h(X) \in R[X]$ such that $\psi^*(X) := \psi(X) - \varphi_\alpha(X)h(X)$ is of degree less than d. Note that $\psi^*(\alpha) = \psi(\alpha) - \varphi_\alpha(\alpha)h(\alpha) = 0$. It follows that $\psi^*(X) = 0$ in $R[X]$ because α is an element of degree d over K. Hence $\psi(X) = \varphi_\alpha(X)h(X) \in R[X]$, which means that $\lambda \in R$.

The implications: (i) $\Rightarrow$ (iii') and (ii) $\Rightarrow$ (i) can be proved by the similar way to the above argument. ∎

PROPOSITION 1.7. *Assume that* $\operatorname{grade}(I_{[\alpha]} + I_{[\alpha^{-1}]}) > 1$. *Then* $R[\alpha, \alpha^{-1}] \cap K = R$ *if and only if* $R[\alpha] \cap K = R$ *and* $R[\alpha^{-1}] \cap K = R$ *(i.e., both* α *and* α^{-1} *are exclusive over* R*).*

PROOF. ($\Leftarrow$): Take $p \in \mathrm{Dp}_1(R)$. Then $I_{[\alpha]} \not\subseteq p$ or $I_{[\alpha^{-1}]} \not\subseteq p$. By the assumption $\operatorname{grade}(I_{[\alpha]} + I_{[\alpha^{-1}]}) > 1$.

Assume that $I_{[\alpha]} \not\subseteq p$. Then $(I_{[\alpha]})_p = R_p$, so that $\varphi_\alpha(X) \in R_p[X]$. Since $\varphi_\alpha(\alpha) = \alpha^d + \eta_1\alpha^{d-1} + \cdots + \eta_d = 0$, we have $\alpha = -(\eta_1 + \eta_2\alpha^{-1} + \cdots + \eta_d\alpha^{1-d}) \in R_p[\alpha^{-1}]$. Thus $R_p[\alpha] \subseteq R_p[\alpha^{-1}]$ and hence $R_p[\alpha, \alpha^{-1}] = R_p[\alpha^{-1}]$. So we have $R_p[\alpha, \alpha^{-1}] \cap K = R_p$ by the assumption α^{-1} is exclusive over R_p (*cf.* Remark 1.5). Next assume that $I_{[\alpha^{-1}]} \not\subseteq p$. Then in the similar way to the above case, we have $R_p[\alpha, \alpha^{-1}] \cap K = R_p$ by the assumption α is exclusive over R_p (*cf.* Remark 1.5).

Thus we have $R \subseteq R[\alpha, \alpha^{-1}] \cap K \subseteq \bigcap_{p \in \mathrm{Dp}_1(R)} R_p[\alpha, \alpha^{-1}] \cap K = \bigcap_{p \in \mathrm{Dp}_1(R)} R_p = R$.

The implication ($\Rightarrow$) follows from the inclusions $R \subseteq R[\alpha^{\pm 1}] \cap K \subseteq R[\alpha, \alpha^{-1}] \cap K = R$.

2. Contractions of principal ideals of $R[\alpha] \cap R[\alpha^{-1}]$. In this section, we treat the problems (ii) and (iii) mentioned as above. Recall that $B_\alpha := R[\alpha] \cap R[\alpha^{-1}]$.

PROPOSITION 2.1. *Let α be an algebraic element of degree d over R and let β be a non-zero element in B_α. Then*

$$\beta B_\alpha \cap R = B_\alpha :_R \beta^{-1}.$$

PROOF. Take $a \in B_\alpha \cap R$. Then $a = \beta y$ with some $y \in B_\alpha$. Hence $y \in a\beta^{-1} \in B_\alpha$. Thus $a \in B_\alpha : R\beta^{-1}$.

Conversely, take $x \in R$ such that $x\beta^{-1} \in B_\alpha$. Then $x \in \beta B_\alpha \cap R$. ∎

REMARK 2.2. Under the same notation as in Proposition 2.1, we showed that $\gamma R[\alpha] \cap R = R[\alpha] :_R \gamma^{-1}$ for a non-zero element $\gamma \in R[\alpha]$ (*cf.* [OY3, Lemma 7].

Now recall the following structure theorem of the subring $B_\alpha := R[\alpha] \cap R[\alpha^{-1}]$.

LEMMA 2.3 (*cf.* [KY3, THEOREM 1]). *Put $\zeta_i := \alpha^i + \eta_1\alpha^{i-1} + \cdots + \eta_i$ for $1 \leq i \leq d-1$ and $\eta_i \in K$. Assume that α is an anti-integral element of degree d over R. Then $B_\alpha = R \oplus I_{[\alpha]}\zeta_1 \oplus \cdots \oplus I_{[\alpha]}\zeta_{d-1}$.*

COROLLARY 2.3.1. *Assume that α is an anti-integral element of degree d over R. Then $B_\alpha \cap K = R$ (i.e., B_α is exclusive over R).*

REMARK 2.4. Take $\beta \in R[\alpha]$. Note that $K(\alpha) = K[\alpha] = K + K\alpha + \cdots + K\alpha^{d-1}$, and put $\beta^{-1} := \lambda_0 + \lambda_1\alpha + \cdots + \lambda_{d-1}\alpha^{d-1}$ with $\lambda_i \in K$. In [OY3, Theorem 8], we showed that if $\mathrm{grade}(\bigcap_{i=0}^{d-1} I_{\lambda_i} + I_{[\alpha]}) > 1$ (of course, $I_{[\alpha]} = R$ is sufficient, *cf.* [OY3, Proposition 4]), then $\beta R[\alpha] \cap R = \bigcap_{i=0}^{d-1} I_{\lambda_i}$.

The following theorem is similar to the known result (*cf.* Remark 2.4). But as for the extension B_α of R, we need not require the assumption such as $\mathrm{grade}(\bigcap_{i=0}^{d-1} I_{\lambda_i} + I_{[\alpha]}) > 1$.

Let ζ_i be the same as in Lemma 2.3. Then $K[\alpha] = K(\alpha) = K(\zeta_1, \ldots, \zeta_{d-1}) = K[\zeta_1, \ldots, \zeta_{d-1}]$.

THEOREM 2.5. *Assume that α is an anti-integral element of degree d over R. Let β be a non-zero element in B_α and let $\beta^{-1} := \mu_0 + \mu_1\zeta_1 + \cdots + \mu_{d-1}\zeta_{d-1}$ with $\mu_i \in K$. Then*

$$\beta B_\alpha \cap R = I_{\mu_0} \cap \bigcap_{i=1}^{d-1} (I_{[\alpha]} :_R \mu_i).$$

PROOF. Since α is super-primitive over R, $B_\alpha = R \oplus I_{[\alpha]}\zeta_1 + \oplus \cdots \oplus I_{[\alpha]}\zeta_{d-1}$ by Lemma 2.3. Hence using Proposition 2.1, we have $\beta B_\alpha \cap R = B_\alpha : R\beta^{-1} = (R \oplus I_{[\alpha]}\zeta_1 + \oplus \cdots \oplus I_{[\alpha]}\zeta_{d-1}) :_R (\mu_0 + \mu_1\zeta_1 + \cdots + \mu_{d-1}\zeta_{d-1}) = I_{\mu_0} \cap \bigcap_{i=1}^{d-1} (I_{[\alpha]} :_R \mu_i)$. ∎

3. Grade of ideals in $R[\alpha] \cap R[\alpha^{-1}]$. The following result is concerned with Problem (iii).

LEMMA 3.1. *Assume that α is an anti-integral element of degree d over R. If $a, b \in R$ is a regular sequence on R, then so is on B_α.*

PROOF. Take $P \in \mathrm{Dp}_1(B_\alpha)$ such that $a, b \in P$. Then $a, b \in p := P \cap R$. But it is known that $p \in \mathrm{Dp}_1(R)$ (*cf.* Lemma 2.3), which is a contradiction. ∎

PROPOSITION 3.2. *Assume that α is an anti-integral element of degree d over R. Let $\beta_1, \beta_2 \in B_\alpha$.*

Assume also that $\mathrm{grade}\big(I_{[\alpha]} + (\beta_i B_\alpha \cap R)\big) > 1$ *for* $i = 1, 2$. *If* $\mathrm{grade}\big((\beta_1 B_\alpha \cap R) + (\beta_2 B_\alpha \cap R)\big) > 1$, *then β_1, β_2 form a regular sequence on B_α.*

PROOF. Take $P \in \mathrm{Dp}_1(B_\alpha)$ such that $\beta_1, \beta_2 \in P$. Put $p := P \cap R$. Then $p \in \mathrm{Dp}_1(R)$ by Lemma 3.1. By construction, we have $p \supseteq (\beta_1 B_\alpha \cap R) + (\beta_2 B_\alpha \cap R)$, which shows that $\mathrm{grade}(p) \geq 2$, a contradiction. ∎

THEOREM 3.3. *Assume that α is an anti-integral element of degree d over R. Let $\beta_1, \ldots, \beta_k \in B_\alpha$.*

Assume further that $J_{[\alpha]} + \sum_{i=1}^{k} (\beta_i B_\alpha \cap R) = R$. *If* $\mathrm{grade}\big(\sum_{i=1}^{k} (\beta_i B_\alpha \cap R)\big) \geq n$, *then* $\mathrm{grade}\big((\beta_1, \ldots, \beta_k) B_\alpha\big) \geq n$.

PROOF. Take $P \in \mathrm{Spec}(B_\alpha)$ such that $(\beta_1, \ldots, \beta_k) \subseteq P$. Put $p := P \cap R$. Then $J_{[\alpha]} \not\subseteq P$ by the assumption. So $R_p[\alpha]$ is flat over R_p (*cf.* [OSaY]) and hence $(B_\alpha)_p$ is flat over R_p, for $(J_{[\alpha]})_p = R_p$ (*cf.* [OSaY]) implies that $(I_{[\alpha]})_p$ is an invertible ideal of R_p and $(B_\alpha)_p = R_p \oplus (I_{[\alpha]})_p\zeta_1 \oplus \cdots \oplus (I_{[\alpha]})_p\zeta_{d-1}$ by Lemma 2.3. Thus the flatness guarantees that the grade does not decrease. ∎

REFERENCES

[KY] M. Kanemitsu and K. Yoshida, *Some properties of extensions $R[\alpha] \cap R[\alpha^{-1}]$ over Noetherian domains*. Comm. Algebra **23**(1995), 4501–4507.

[M] H. Matsumura, *Commutative Algebra*. 2nd ed., Benjamin, New York, 1980.

[OSaY] S. Oda, J. Sato and K. Yoshida, *High degree anti-integral extensions of Noetherian domains*. Osaka J. Math. **30**(1993), 119–135.

[**OY1**] S. Oda and K. Yoshida, *Anti-integral extensions of Noetherian domains.* Kobe J. Math. **5**(1988), 43–56.

[**OY2**] ———, *Remarks on an exclusive extension generated by a super-primitive element.* Osaka J. Math. **32**(1995), 495–499.

[**OY3**] ———, *On contractions of principal ideals and denominator ideals in simple algebraic extensions.* Algebra Colloq. **5**(1998), 355–360.

[**YSaO**] K. Yoshida, J. Sato and S. Oda, *On exclusive extensions of Noetherian domains.* Bull. Okayama Univ. of Sci. **29**(1994), 21–26.

Matsusaka Commercial High School
Toyohara
Matsusaka
Mie 515-0205
Japan

Department of Applied Mathematics
Okayama University of Science
Ridai-cho
Okayama 700-0005
Japan

C. R. Math. Rep. Acad. Sci. Canada Vol. 21 (2), 1999 pp. 46–52

DIHEDRAL EXTENSIONS IN CHARACTERISTIC 0

ARNE LEDET

Presented by M. Ram Murty, FRSC

ABSTRACT. We describe a systematic (algorithmic) method for finding generic polynomials for dihedral extensions of odd degree in characteristic 0.

RÉSUMÉ. On donne une méthode systematique (algorithmique) pour trouver des polynomes génériques pour des extensions dihédrales de degré impair en charactéristique 0.

Introduction. One of the problems of Inverse Galois Theory is the following: Given a field K and a finite group G, what do G-extensions of K look like? And how few assumptions about K must we make in order for the description to work?

One approach to this problem is the construction of generic polynomials:

DEFINITION.. Let K be a field and G a finite group, and let $\mathbf{t} = (t_1, \ldots, t_n)$ and X be indeterminates over K. A polynomial $F(\mathbf{t}, X) \in K(\mathbf{t})[X]$ is called *generic* for G-extensions over K, if it has the following properties:

(1) The splitting field of $F(\mathbf{t}, X)$ over $K(\mathbf{t})$ is a G-extension.

(2) If L/K is a field extension, any G-extension of L is obtained as the splitting field of $F(\mathbf{a}, X)$ for a suitable $\mathbf{a} \in L^n$.

Thus, if $F(\mathbf{t}, X)$ is generic for G-extensions over K, the splitting field of $F(\mathbf{t}, X)$ over $K(\mathbf{t})$ is a sort of 'once and for all' description of G-extensions over fields $\supseteq K$.

A more precise way of giving such descriptions is generic extensions, introduced by Saltman in [Sa, 1982]:

DEFINITION. Let K be a field and G a finite group. A Galois extension S/R of commutative rings with group G is called a *generic G-extension* over K, if it has the following properties:

(1) $R = K[\mathbf{t}, 1/t]$ for indeterminates $\mathbf{t} = (t_1, \ldots, t_n)$ and some element $t \in K[\mathbf{t}] \setminus 0$.

(2) If L/K is a field extension, any G-extension T/L is obtained as a specialisation of S/R. (*I.e.*, $T/L \simeq S \otimes_\varphi L/L$ as G-extensions for some K-algebra homomorphism $\varphi\colon R \to L$.)

Received by the editors 7 December, 1998.

This work was supported by a Queen's University Advisory Research Committee Postdoctoral Fellowship.

AMS subject classification: 12F12.

Note that a generic extension contains more information than a generic polynomial, since the G-extension T of L in the definition is not required to be a field, but simply a Galois algebra.

In this paper, we will construct generic polynomials for dihedral groups of odd degree n over the field $\mathbb{Q}$ of rational numbers. This extends the result of Hashimoto and Miyake in [H&M], where such a polynomial is constructed over the field $\mathbb{Q}(\cos\frac{2\pi}{n})$. For our purposes, the definition of a dihedral group is

DEFINITION. Let $n \geq 3$ be a natural number. Then the *dihedral group* of degree n is the group D_n with generators σ and τ and relations $\sigma^n = \tau^2 = 1$ and $\tau\sigma = \sigma^{-1}\tau$.

More generally, if A is an abelian group, the dihedral group D_A is the semi-direct product $A \rtimes C_2$, where the cyclic group C_2 acts on A by inversion. If A is the direct product of cyclic groups of orders $q_1, \ldots, q_r$, we write $D_{q_1 \times \cdots \times q_r}$ instead of D_A.

The construction of generic polynomials will be via generic extensions. Generic extensions give rise to generic polynomials in the following way:

Let S/R be a generic G-extension over a field K. For $s \in S$, we define $\mathrm{Min}(s,R) = \prod_{s' \in Gs}(X - s')$, *i.e.*, $\mathrm{Min}(s,R)$ is the product of the distinct conjugates of $X - s$ under G-action. It is clear that $\mathrm{Min}(s,R) \in R[X]$, and we claim that $\mathrm{Min}(s,R)$ is separable in $\mathbb{K}[X]$, where $\mathbb{K} = K(\mathbf{t})$ is the quotient field of R, *i.e.*, that the irreducible factors of $\mathrm{Min}(s,R)$ have no multiple roots: Let $\mathfrak{m}$ be a maximal ideal in the Galois algebra $S \otimes_R \mathbb{K}$, and let $\mathbb{L} = S \otimes_R \mathbb{K}/\mathfrak{m}$. Then $\mathbb{L}$ is a simple component of $S \otimes_R \mathbb{K}$, and hence a Galois field extension of $\mathbb{K}$ with Galois group $H \subseteq G$. Also, $\mathrm{Min}(s,R)$ splits in linear factors over $\mathbb{L}$, since it does over S. Hence, $\mathrm{Min}(s,R)$ is separable.

Now, by [D&I, Ch. III Prop. 1.2], S is a finitely generated projective R-module. Let $s_1, \ldots, s_m$ generate S over R, and let $f(X) = \prod_i \mathrm{Min}(s_i,R)$. Then $f(X)$ is a monic separable polynomial in $\mathbb{K}[X]$. The splitting field of $f(X)$ over $\mathbb{K}$ is the simple component $\mathbb{L}$ from above.

Next, let M/L be a G-extension of fields with $L \supseteq K$. Then there is a specialisation $\varphi\colon R \to L$, such that $S \otimes_\varphi L \simeq M$. It follows that M is the splitting field of $\varphi(f)(X) \in L[X]$. It also follows that $H = G$, since the splitting field over L of $\varphi(f)(X)$ has a Galois group of order $\leq |H|$.

In order to carry out the construction of generic polynomials of dihedral groups, we will first need to recall Saltman's [Sa] results about generic extensions for cyclic groups. This is done in Section 1 below. Section 2 then gives the actual construction for the dihedral group D_q, when q is an odd prime power, and Section 3 gives various immediate consequences.

1. Cyclic extensions. In [Sa, Section 2], Saltman constructs generic C_q-extensions for odd prime powers $q = p^n$ under the assumption that the base field

has characteristic $\neq p$. In particular, the construction works over $\mathbb{Q}$. We quote the construction here, omitting the proof:

Let $d = p^{n-1}(p-1)$ and let $e \in \mathbb{Z}$ be a generator for $\mathbb{Z}/p^2$. Then e generates $\mathbb{Z}/p^m$ for all m, and in particular we have $p \nmid (e^d - 1)/q$, as required by Saltman.

We let $\mathbb{Q}_q = \mathbb{Q}(\zeta)$, where $\zeta = \exp(2\pi i/q)$, and let $\kappa \in \mathrm{Gal}(\mathbb{Q}_q/\mathbb{Q})$ be given by $\kappa\zeta = \zeta^e$. κ is then a generator for $\mathrm{Gal}(\mathbb{Q}_q/\mathbb{Q})$.

Now, we define a map Φ by $\Phi(x) = x^{e^{d-1}} \kappa x^{e^{d-2}} \cdots \kappa^{d-1} x$. Whenever κ extends from $\mathbb{Q}_q$ to a $\mathbb{Q}_q$-algebra R', Φ is defined on R'.

Let $R' = \mathbb{Q}_q[x_1, \ldots, x_d, 1/x]$, where $x_1, \ldots, x_d$ are indeterminates and $x = x_1 \cdots x_d$. We extend κ to R' by $\kappa x_i = x_{i+1}$ ($\kappa x_d = x_1$), and let $S' = R'[\theta]$, where $\theta^n = \Phi(x_1)$. Then κ extends to S' by $\kappa\theta = x_1^{-(e^d-1)/q}\theta^e$, and this extension has order d. Also, S'/R' is a C_q-extension, and the Galois group is generated by $\sigma\colon \theta \mapsto \zeta\theta$.

Since κ and σ commutes, S/R is a C_q-extension, when $S = S'^{\kappa}$ and $R = R'^{\kappa}$. Furthermore, letting $y_i = \sum_{j=1}^{d} \zeta^{(i-1)e^{j-1}} x_j$, we get that $y_1, \ldots, y_d$ is a basis for the $\mathbb{Q}_q$-vector space of monomials in $\mathbb{Q}_q[x_1, \ldots, x_d]$, and that the y_i's are κ-invariant. In other words, $R = \mathbb{Q}[y_1, \ldots, y_d, 1/x]$ and $x \in \mathbb{Q}[y_1, \ldots, y_d]$.

We now have

THEOREM 1. *S/R is a generic C_q-extension over $\mathbb{Q}$.*

For use in studying D_q-extensions, we wish to construct a normal basis for S/R. By a *normal basis* we mean an R-basis for S of the form $(\beta, \sigma\beta, \ldots, \sigma^{q-1}\beta)$ for an $\beta \in S$. Clearly, we can obtain such a basis by producing a κ-invariant normal basis for S'/R'. (For a comprehensive treatment of normal bases of cyclic extensions, including descent, see [Gr, Ch. I].)

First, we note that an element $\beta = \sum_{i=0}^{q-1} a_i\theta^i$ generates a normal basis for S'/R' if $a_i \in R'^*$ for all i. This is obvious, since the matrix transforming the basis $(1, \theta, \ldots, \theta^{q-1})$ into $(\sigma\beta)_{\sigma \in C_q}$ is invertible.

We claim that the a_i's can be chosen such that $\beta \in S$:

Clearly, we can replace the basis $(1, \theta, \ldots, \theta^{q-1})$ by a basis made up of (1), $(\theta, \kappa\theta, \ldots, \kappa^{d-1}\theta)$, $(\theta^p, \kappa\theta^p, \ldots, \kappa^{d/p-1}\theta^p)$, $\ldots$, $(\theta^{p^{n-1}}, \kappa\theta^{p^{n-1}}, \ldots, \kappa^{d/p^{n-1}-1}\theta^{p^{n-1}})$, since we get all the exponents $0, 1, \ldots, q-1$ of θ. Representing β in this basis, the condition $\kappa\beta = \beta$ translates as

$$\begin{aligned}\beta = {} & a_0 + a_1\theta + \kappa a_1 \kappa\theta + \cdots + \kappa^{d-1} a_1 \kappa^{d-1}\theta \\ & + a_p\theta^p + \kappa a_p \kappa\theta^p + \cdots + \kappa^{d/p-1} a_p \kappa^{d/p-1}\theta^p + \cdots\end{aligned}$$

where $a_0 \in R$, and $\kappa^{d/p^i} a_{p^i} \kappa^{d/p^i} \theta^{p^i} = a_{p^i}\theta^{p^i}$ for $i = 0, \ldots, n-1$.

The element $\kappa^{d/p^i}\theta^{p^i}/\theta^{p^i}$ is in R'^* (since e has order d/p^i modulo q/p^i) and has norm 1 with respect to κ^{d/p^i}. Writing $\kappa^{d/p^i}\theta^{p^i}/\theta^{p^i} = x_1^{e_1} \cdots x_d^{e_d}$, this simply

means that $e_j + e_{d/p^i+j} + \cdots + e_{(p^i-1)d/p^i+j} = 0$ for all $j = 1, \ldots, d/p^i$, and we see that we can let

$$a_{p^i} = \prod_{j=1}^{d/p^i} \prod_{k=0}^{p^i-1} x_{kd/p^i+j}^{-(e_j+e_{d/p^i+j}+\cdots+e_{kd/p^i+j})}.$$

Of course, we can modify a_{p^i} by a power of $x_1 \cdots x_d$ if we want to. (For instance, to make β integral over $\mathbb{Q}[y_1, \ldots, y_d]$.)

2. **Dihedral extensions.** In [Sa, Section 3], Saltman proves various results concerning generic extensions for semi-direct products, most notably wreath products. These results implies the existence of generic extensions for dihedral groups of odd degree over any field. In this section we adapt these results to get generic extensions over $\mathbb{Q}$ particularly suited for finding generic polynomials.

We consider the dihedral group D_q, where $q = p^n$ is an odd prime power, and start by making the following observation: $C_q \times D_q \simeq C_q \wr C_2$, and the factor C_q on the left corresponds to the C_2-invariant subgroup N of $C_q \times C_q$ on the right. (Here, $C_q \wr C_2$ is the *wreath product* of C_q and C_2, *cf.* [Hu, I., Section 15], *i.e.*, the semi-direct product of $C_q \times C_q$ and C_2 with C_2 acting by swapping the coordinates.) So, whenever we have a $C_q \wr C_2$-extension S/K, we get (canonically) a D_q-subextension S^N/K. On the other hand, if S'/K is a D_q-extension, $S/K = K^q \otimes_K S'/K$ is a $C_q \wr C_2$-extension and $S^N = S'$.

Thus, we can study $C_q \wr C_2$-extensions in the assurance that we will get all D_q-extensions in the process.[1] This is a clear advantage, since $C_q \wr C_2$ is easier to handle than D_q.

Now, assume the following: S/R is a generic C_q-extension over $\mathbb{Q}$, σ generates the Galois group, and $R = \mathbb{Q}[\mathbf{y}, 1/y]$, where $\mathbf{y} = (y_1, \ldots, y_d)$ are indeterminates and $y \in \mathbb{Q}[\mathbf{y}] \setminus 0$.

We let $R_1 = \mathbb{Q}[\mathbf{s}, \mathbf{t}, u, 1/su]$, where $\mathbf{s} = (s_1, \ldots, s_d)$, $\mathbf{t} = (t_1, \ldots, t_d)$ and s are indeterminates, and $s = y(s_1 + t_1\sqrt{u}, \ldots, s_d + t_d\sqrt{u})y(s_1 - t_1\sqrt{u}, \ldots, s_d - t_d\sqrt{u}) \in \mathbb{Q}[\mathbf{s}, \mathbf{t}, u]$. Also, we let $R_2 = \mathbb{Q}[\mathbf{s}, \mathbf{t}, \sqrt{u}, 1/su]$. Then R_2/R_1 is a generic C_2-extension over $\mathbb{Q}$. We denote the generator for the Galois group of R_2/R_1 by κ.

Next, we define homomorphisms $\varphi_1, \varphi_2 \colon R \to R_2$ by $\varphi_1(y_i) = s_i + t_i\sqrt{u}$ and $\varphi_2(y_i) = s_i + t_i\sqrt{u}$. This gives us specialisations $S_1 = S \otimes_{\varphi_1} R_2$ and $S_2 = S \otimes_{\varphi_2} R_2$, that will be C_q-extensions of R_2 with generators σ_1 and σ_2. We can then extend κ to an R_1-isomorphism $S_1 \simeq S_2$ by $\kappa(s \otimes r_2) = s \otimes \kappa r_2$. This works both ways to give us $\kappa^2 = 1$ and $\kappa\sigma_1 = \sigma_2\kappa$. Thus, we have κ acting on $T = S_1 \otimes_{R_2} S_2$. It is almost obvious that T/R_1 is a $C_q \wr C_2$-extension, and we claim that it is in fact generic:

[1] Provided, of course, that we look at Galois *algebras*, and not just fields.

Let U/K be a $C_q \wr C_2$-extension in characteristic 0. Then we have $U = T_1 \otimes_L T_2$, where L/K is the quadratic subextension, and T_1/L and T_2/L are conjugate C_q-extensions. Now, $L = K[\bar{u}]$ for some $\bar{u}^2 = a \in K^*$, and T_1/L is obtained by specialising S/R with respect to a map $y_i \mapsto a_i + b_i\bar{u}$. The map $\psi\colon R_2 \to L$, given by $\sqrt{u} \mapsto \bar{u}$, $s_i \mapsto a_i$ and $t_i \mapsto b_i$, will then give T_1/L by specialisation as well. Also, since $\kappa S_1 = S_2$ and $\kappa T_1 = T_2$, the same specialisation give us T_2/L from S_2/R_2, and hence U/L from T/R_2. Letting $\varphi = \psi|_{R_1}$, this means that $T \otimes_\varphi K \simeq U$.

It follows that T^N/R_1 is generic for D_q-extensions over $\mathbb{Q}$.

Now, generic C_q-extensions were described in Section 1, where it was established that they can be constructed to have normal bases: There is an element $\beta \in S$ such that $\beta, \sigma\beta, \dots, \sigma^{q-1}\beta$ is a basis for S/R. Looking at T/R_1 above, this means that there are elements β_1 and $\beta_2 = \kappa\beta_1$ in T, such that $\beta_i, \sigma_i\beta_i, \dots, \sigma_i^{q-1}\beta_i$ is a basis for S_i/R_2, and hence such that $(\sigma_1^i\beta_1 \otimes \sigma_2^j\beta_2)_{0 \le i,j < q}$ is a basis for T/R_2.

The trace $\mathrm{Tr}_{T/T^N}\colon T \to T^N$ is surjective and R_2-linear, and so the traces of the elements $\sigma_1^i\beta_1 \otimes \sigma_2^j\beta_2$ generate T^N over R_2. Since there are only q distinct traces $\alpha_i = \sum_{j=0}^{q-1} \sigma_1^j\beta_1 \otimes \sigma_2^{i+j}\beta_2$, these elements form a basis for T^N/R_2. Also, as they are conjugate, it is a normal basis.

Let $f(\mathbf{s}, \mathbf{t}, u, X) = \prod_{i=0}^{q-1}(X - \alpha_i)$. Then $f \in \mathbb{Q}(\mathbf{s}, \mathbf{t}, u)[X]$, since α_0 is κ-invariant. As T^N/R_1 is a generic D_q-extension, we have the following: For every D_q-extension L/K in characteristic 0, there is a specialisation of f over K with splitting field L over the quadratic subextension of L/K. This immediately implies that L is in fact the splitting field over K of the specialised polynomial, and we conclude that f is generic for D_q-extensions over $\mathbb{Q}$.

PROPOSITION 2. *A generic polynomial for D_q-extensions over $\mathbb{Q}$ exists and can be explicitly constructed.*

In fact, assume that an element $\beta = \sum_{i=0}^{q-1} a_i\theta^i$ generating a normal basis for S/R has been found as described in Section 1, where a_i is a rational monomial in $x_1, \dots, x_d$. Then the construction is follows:

As above, $q = p^n$, $d = p^{n-1}(p-1)$ and e generates $\mathbb{Z}/q$. We introduce indeterminates u, $\mathbf{s} = (s_1, \dots, s_d)$ and $\mathbf{t} = (t_1, \dots, t_d)$. In $\mathbb{Q}(\mathbf{s}, \mathbf{t}, \sqrt{u})$ we let new 'indeterminates' $\mathbf{x}_1 = (x_{11}, \dots, x_{1d})$ and $\mathbf{x}_2 = (x_{21}, \dots, x_{2d})$ be given by

$$s_j + t_j\sqrt{u} = \sum_{i=1}^{d} \zeta^{(i-1)e^{j-1}} x_{1i} \quad \text{and} \quad s_j - t_j\sqrt{u} = \sum_{i=1}^{d} \zeta^{(i-1)e^{j-1}} x_{2i},$$

where $\zeta = \exp(2\pi i/q)$. Next, we let

$$\theta_1 = \sqrt[q]{x_{11}^{e^{d-1}} x_{12}^{e^{d-2}} \cdots x_{1d}} \quad \text{and} \quad \theta_2 = \sqrt[q]{x_{21}^{e^{d-1}} x_{22}^{e^{d-2}} \cdots x_{2d}}.$$

With

$$\beta_1 = a_0(\mathbf{x}_1) + a_1(\mathbf{x}_1)\theta_1 + \cdots + a_{q-1}(\mathbf{x}_1)\theta_1^{q-1},$$
$$\beta_2 = a_0(\mathbf{x}_2) + a_1(\mathbf{x}_2)\theta_2 + \cdots + a_{q-1}(\mathbf{x}_2)\theta_2^{q-1},$$

the generic polynomial is

$$f(\mathbf{s},\mathbf{t},u,X) = \prod_{i=0}^{q-1}\Big(X - \sum_{j=0}^{q-1}\sigma_1^i\beta_1\sigma_2^{i+j}\beta_2\Big),$$

where σ_1 and σ_2 are given by $\sigma_1\theta_1 = \zeta\theta_1$, $\sigma_1\theta_2 = \theta_2$, $\sigma_2\theta_1 = \theta_1$ and $\sigma_2\theta_2 = \zeta\theta_2$.

EXAMPLE. Look at the simplest case, $q = 3$. Then $d = e = 2$, and $\Phi(x_1) = x_1^2x_2$. We can replace this by $x_2/x_1 = x_1^2x_2/(x_1)^3$ to get $\theta^3 = x_2/x_1$ and $\kappa\theta = 1/\theta$. Thus, we can let $\beta = 1 + \theta + 1/\theta$. The construction (conveniently carried out by computer) then gives us

$$f(s_1,s_2,t_1,t_2,u,X) = X^3 - 9X^2 + \frac{324(s_1t_2 - s_2t_1)^2u}{S^2 - T^2u} \in \mathbb{Q}(s_1,s_2,t_1,t_2,u)[X],$$

as a generic polynomial for D_3-extensions over $\mathbb{Q}$, where

$$S = s_1^2 + s_1s_2 + s_2^2 + u(t_1^2 + t_1t_2 + t_2^2),$$
$$T = 2s_1t_1 + s_1t_2 + s_2t_1 + 2s_2t_2.$$

From this we can easily deduce that $X^3 + X^2 + t \in \mathbb{Q}(t)[X]$ is generic as well, but f has the advantage that it allows us to 'control' the quadratic subextension, since this is given by u.

3. Generalised dihedral extensions. The construction in the preceding section immediately gives us various additional results:

PROPOSITION 3. *Let $q_1,\ldots,q_r$ be powers of odd (not necessarily distinct) primes. Then there is a generic polynomial for $D_{q_1\times\cdots\times q_r}$-extensions over $\mathbb{Q}$.*

PROOF. Let $f_i(\mathbf{s}_i,\mathbf{t}_i,u,X)$ be generic for D_{q_i} as above. Since a $D_{q_1\times\cdots\times q_r}$-extension is the composite of D_{q_1}-, ..., D_{q_r}-extensions with the same quadratic subextension, it is clear that

$$f_1(\mathbf{s}_1,\mathbf{t}_1,u,X)\cdots f_r(\mathbf{s}_r,\mathbf{t}_r,u,X) \in \mathbb{Q}(\mathbf{s}_1,\ldots,\mathbf{s}_r,\mathbf{t}_1,\ldots,\mathbf{t}_r,u)[X]$$

is generic for $D_{q_1\times\cdots\times q_r}$-extensions over $\mathbb{Q}$. ∎

In particular, this allows for construction of generic D_n-polynomials for all odd numbers $n \geq 3$, since $D_{q\times q'} = D_{qq'}$ when q and q' are mutually prime. More generally, we have

PROPOSITION 4. *Let $n \geq 3$ be an integer not divisible by 16. Then there is a generic polynomial for D_n-extensions over $\mathbb{Q}$.*

PROOF. If n is odd, it follows from Proposition 3 that there is a generic polynomial $f_n(\mathbf{s},\mathbf{t},u,X)$, where specialisation of u gives the quadratic subextension.

If $n = 2m$, m odd, we have $D_n = C_2 \times D_m$, and so $f_m(\mathbf{s},\mathbf{t},u,X)(X^2 - v)$ is generic.

If $n = 4m$, m odd: $X^4 - 2stX^2 + s^2t(t-1) \in \mathbb{Q}(s,t)[X]$ is generic for D_4 over $\mathbb{Q}$ (and in fact in characteristic $\neq 2$) as an easy consequence of [Ki, Th. 5], and the quadratic subextension (*i.e.*, the one we are interested in) is given by specialisation of $t-1$. Thus, $f_m(\mathbf{s},\mathbf{t},t-1,X)\big(X^4 - 2stX^2 + s^2t(t-1)\big)$ is generic for D_n, if $f_m(\mathbf{s},\mathbf{t},u,X)$ is generic for D_m as above.

If $n = 8m$, m odd: By [Le, Th. 2], there is a generic polynomial $G(x,y,z,r,s,X)$ for D_8-extensions in characteristic $\neq 2$, and the interesting quadratic subextension is given by specialisation of $(1-2y^2)/(1+x^2-2z^2) - 1$. Thus,

$$f_m\big(\mathbf{s},\mathbf{t},(1-2y^2)/(1+x^2-2z^2)-1,X\big)G(x,y,z,r,s,X) \in \mathbb{Q}(\mathbf{s},\mathbf{t},r,s,x,y,z)[X]$$

is generic for D_n, when $f_m(\mathbf{s},\mathbf{t},u,X)$ is as above. ∎

Thus, the example in Section 2 makes it possible to describe generic polynomials for $D_{3\times 3}$-, D_6-, D_{12}- and D_{24}-extensions, among others, starting from the D_3-polynomial given.

REMARK. Generic D_8-extensions are constructed by Black in [Bl, Th. 4.6].

REFERENCES

[Bl] E. V. Black, *Deformations of Dihedral 2-Group Extensions of Fields.* Trans. Amer. Math. Soc. To appear.

[D&I] F. DeMeyer and E. Ingraham, *Separable Algebras over Commutative Rings.* Lecture Notes in Math. **181**, Springer-Verlag, 1971.

[Gr] C. Greither, *Cyclic Galois Extensions of Commutative Rings.* Lecture Notes in Math. **1534**, Springer-Verlag, 1992.

[H&M] K. Hashimoto and K. Miyake, *Inverse Galois Problem for Dihedral Groups.* Developments in Mathematics, Kluwer. To appear.

[Hu] B. Huppert, *Endliche Gruppen I.* Grundlehren Math. Wiss. **134**, Springer-Verlag, 1967.

[Ki] I. Kiming, *Explicit classifications of some 2-extensions of a field of characteristic different from 2.* Canad. J. Math. **42**(1990), 825–855.

[Le] A. Ledet, *Generic polynomials for quasi-dihedral, dihedral and modular extensions of order 16.* Preprint, 1998.

[Sa] D. Saltman, *Generic Galois Extensions and Problems in Field Theory.* Adv. Math. **43**(1982), 250–283.

Department of Mathematics and Statistics
Queen's University
Kingston, Ontario K7L 3N6
email: ledet@mast.queensu.ca

C. R. Math. Rep. Acad. Sci. Canada Vol. **21** (1), 1999 pp. 53–59

INÉGALITÉS EXPLICITES POUR $\psi(X)$, $\theta(X)$, $\pi(X)$ ET LES NOMBRES PREMIERS

PIERRE DUSART

Présenté par P. Ribenboim, FRSC

ABSTRACT. In default of a proof of the Riemann hypothesis, the best estimates for $\psi(x)$ and $\theta(x)$ and hence of $\pi(x)$, p_k and other functions of the primes, depend on the current state of knowledge of the zeros of $\zeta(s)$. With a better knowledge about the zeros of the Riemann ζ function, we can show sharper bounds for $\psi(x)$,$\theta(x)$,$\pi(x)$ and primes p_k.

RÉSUMÉ. Une meilleure connaissance du positionnement des zéros de la fonction ζ de Riemann permet d'obtenir des estimations effectives plus précises des fonctions $\psi(x)$, $\theta(x)$, $\pi(x)$ et des p_k.

On notera $\ln_2 x$ pour $\ln\ln x$.

1. Fonctions de Chebyshev. Nous nous intéressons aux fonctions de $\psi(x)$ et $\theta(x)$ définies par

$$\theta(x) = \sum_{p \le x} \ln p, \quad \psi(x) = \sum_{\substack{p,\nu \\ p^\nu \le x}} \ln p$$

où les sommes sont respectivement sur tous les p premiers et sur les puissances de premiers p^ν. Le théorème des nombres premiers équivaut à dire que

$$\psi(x) = x + o(x), \quad x \to +\infty.$$

De manière analogue, pour tout $\varepsilon > 0$, il existe $x_0 = x_0(\varepsilon)$ tel que

$$|\psi(x) - x| < \varepsilon x \quad \text{pour } x \ge x_0.$$

Le but de cet article est de donner des estimations explicites de $\theta(x)$ et $\psi(x)$. Cet article s'appuie sur des résultats déjà connus : les plus importants travaux sur les résultats effectifs ont été fournis par Rosser et Schoenfeld (*e.g.* [12], [13], [14]), puis complétés par Robin [9], Robin et Massias [5] et Pereira [4].

Les estimations de $\psi(x)$ dans [13] sont basées sur la vérification de l'hypothèse de Riemann pour les 3 502 500 premiers zéros de $\zeta(s)$ dans la bande critique

Reçu par les éditeurs le 7 December, 1998.
Classification (AMS) par sujet : primaire: 11A25, 11L20; secondaire: 11A41, 11B25.

et sur la connaissance d'une région sans zéro de $\zeta(s)$ de même type que celle trouvée originellement par de la Vallée Poussin. Une meilleure connaissance sur les zéros permet une meilleure estimation de $\psi(x)$. Van de Lune *et al.* [15] ont montré que les 1 500 000 000 premiers zéros sont sur la droite critique. Cette vérification induit un gain notable sur les valeurs "moyennes" de x (c'est-à-dire jusqu'à $\exp(4000)$). Nous en avons déduit de nouveaux encadrements pour $\psi(x)$ et pour $\theta(x)$.

$$|\psi(x)-x| \le 0{,}006409\frac{x}{\ln x} \quad \text{pour } x \ge \exp(22),$$
$$|\theta(x)-x| \le 0{,}006788\frac{x}{\ln x} \quad \text{pour } x \ge 10\,544\,111,$$
$$|\theta(x)-x| \le 0{,}2\frac{x}{\ln^2 x} \quad \text{pour } x \ge 3\,594\,641,$$
$$|\theta(x)-x| \le 515\frac{x}{\ln^3 x} \quad \text{pour } x > 1,$$
$$|\theta(x)-x| \le 1717433\frac{x}{\ln^4 x} \quad \text{pour } x > 1.$$

2. **Résultats sur p_k et $\theta(p_k)$.** Soit p_k, le k-ième nombre premier. Cesaro [2] puis Cipolla [3] donnent l'expression du développement asymptotique en 1902 :

$$p_k = k\left\{\ln k + \ln_2 k - 1 + \frac{\ln_2 k - 2}{\ln k} - \frac{\ln_2^2 k - 6\ln_2 k + 11}{2\ln^2 k} + O\left(\left(\frac{\ln_2 k}{\ln k}\right)^3\right)\right\}.$$

Rosser [10] a montré que $p_k \ge k\ln k$ et améliore son résultat avec Schoenfeld [12] en montrant que $p_k \ge k(\ln k + \ln_2 k - 3/2)$. En 1983, Robin [9] réussit à prouver que

$$p_k \ge k(\ln k + \ln_2 k - 1{,}0072629).$$

Massias et Robin [5] sont capables de montrer que

$$p_k \ge k(\ln k + \ln_2 k - 1)$$

pour $k \le \exp(598)$ et $k \ge \exp(1800)$.

A l'aide des encadrements de ψ, on montre que cela est vrai pour tout $k \ge 2$.

Les encadrements pour p_k et $\theta(p_k)$ sont :

(1) $$\theta(p_k) \ge k\left(\ln k + \ln_2 k - 1 + \frac{\ln_2 k - 2{,}0553}{\ln k}\right) \quad \text{pour } k \ge \exp(22),$$

(2) $$\theta(p_k) \le k\left(\ln k + \ln_2 k - 1 + \frac{\ln_2 k - 2}{\ln k}\right) \quad \text{pour } k \ge 198,$$

(3) $$p_k \ge k\,(\ln k + \ln_2 k - 1) \quad \text{pour } k \ge 2,$$

(4) $$p_k \le k\,(\ln k + \ln_2 k - 0{,}9484) \quad \text{pour } k \ge 39017,$$

(5) $$p_k \le k\left(\ln k + \ln_2 k - 1 + \frac{\ln_2 k - 1{,}8}{\ln k}\right) \quad \text{pour } k \ge 27076,$$

(6) $$p_k \ge k\left(\ln k + \ln_2 k - 1 + \frac{\ln_2 k - 2{,}25}{\ln k}\right) \quad \text{pour } k \ge 2.$$

La formule (2) a été démontrée par Robin dans [9].

2.1. *Intervalle contenant au moins un nombre premier.* Nous connaissons déjà le résultat de Schoenfeld [14] montrant que, pour $x \geq 2010759{,}9$, l'intervalle $]x,x + x/16597[$ contient au moins un nombre premier. Nous améliorerons ce résultat (voir aussi [6] pour d'autres intervalles).

PROPOSITION 1. *Pour $k \geq 463$,*

$$p_{k+1} \leq p_k \left(1 + \frac{1}{2\ln^2 p_k}\right).$$

THÉORÈME 1. *Pour tout $x \geq 3275$, il existe un nombre premier p tel que*

$$x < p \leq x\left(1 + \frac{1}{2\ln^2 x}\right).$$

Ce résultat est meilleur que celui de Rosser et Schoenfeld si $x \geq 3 \cdot 10^{39}$.

3. Résultats sur $\pi(x)$. Nous estimons aussi la fonction qui compte le nombre de nombres premiers inférieurs à x :

$$\pi(x) = \sum_{p \leq x} 1.$$

Rappelons que

$$\pi(x) = \frac{x}{\ln x}\left(1 + \frac{1}{\ln x} + \frac{2}{\ln^2 x} + O\left(\frac{1}{\ln^3 x}\right)\right).$$

Les encadrements effectifs pour $\pi(x)$ sont :

$$\frac{x}{\ln x}\left(1 + \frac{1}{\ln x}\right) \underset{x \geq 599}{\leq} \pi(x) \underset{x > 1}{\leq} \frac{x}{\ln x}\left(1 + \frac{1{,}2762}{\ln x}\right).$$

$$\pi(x) \geq \frac{x}{\ln x - 1} \quad \text{pour } x \geq 5393,$$

$$\pi(x) \leq \frac{x}{\ln x - 1{,}1} \quad \text{pour } x \geq 60184,$$

$$\pi(x) \geq \frac{x}{\ln x}\left(1 + \frac{1}{\ln x} + \frac{1{,}8}{\ln^2 x}\right) \quad \text{pour } x \geq 32299,$$

$$\pi(x) \leq \frac{x}{\ln x}\left(1 + \frac{1}{\ln x} + \frac{2{,}51}{\ln^2 x}\right) \quad \text{pour } x \geq 355991.$$

4. Sommes diverses. La proposition suivante (conjecture de Robert Mandl) a été énoncée dans l'article de Rosser et Schoenfeld [13].

PROPOSITION 2. *Pour $n \geq 9$, nous avons*

$$(p_1 + p_2 + \cdots + p_n)/n < \frac{1}{2}p_n.$$

PROPOSITION 3. *Pour $n \geq 2$, nous avons*

$$p_{[\frac{n}{2}]} \leq \frac{1}{n}\sum_{i=1}^{n} p_i.$$

Conclusion : la moyenne des n premiers nombres premiers est comprise entre le terme médian et le dernier terme divisé par deux.

4.1. *Autres inégalités.* Soit γ la constante d'Euler ($\gamma \approx 0{,}5772156649$).

THÉORÈME 2. *Soit $B = \gamma + \sum_p (\ln(1-1/p) + 1/p) \approx 0{,}26149\ 72128\ 47643$. Pour $x > 1$,*

$$\sum_{p \leq x} \frac{1}{p} - \ln_2 x - B \geq -\left(\frac{1}{10\ln^2 x} + \frac{4}{15\ln^3 x}\right).$$

Pour $x \geq 10372$,

$$\sum_{p \leq x} \frac{1}{p} - \ln_2 x - B \leq \frac{1}{10\ln^2 x} + \frac{4}{15\ln^3 x}.$$

THÉORÈME 3. *Soit $E = -\gamma - \sum_{n=2}^{\infty} \sum_p (\ln p)/p^n \approx -1{,}33258\ 22757\ 33221$. Pour $x > 1$,*

$$\sum_{p \leq x} \frac{\ln p}{p} - \ln x - E \geq -\left(\frac{0{,}2}{\ln x} + \frac{0{,}2}{\ln^2 x}\right).$$

Pour $x \geq 2974$,

$$\sum_{p \leq x} \frac{\ln p}{p} - \ln x - E \leq \frac{0{,}2}{\ln x} + \frac{0{,}2}{\ln^2 x}.$$

THÉORÈME 4. *Pour $x > 1$,*

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) < \frac{e^{-\gamma}}{\ln x}\left(1 + \frac{0{,}2}{\ln^2 x}\right)$$

et

$$\prod_{p \leq x} \frac{\ln p}{p} > e^{\gamma} \ln x \left(1 - \frac{0{,}2}{\ln^2 x}\right).$$

Pour $x \geq 2973$,

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) > \frac{e^{-\gamma}}{\ln x}\left(1 - \frac{0{,}2}{\ln^2 x}\right)$$

et

$$\prod_{p\le x}\frac{\ln p}{p} < e^{\gamma}\ln x\left(1+\frac{0{,}2}{\ln^2 x}\right).$$

THÉORÈME 5. *Pour* $x > 1$,

$$\pi(2x)-\pi(x) \le \frac{x}{\ln x}.$$

Pour $x \ge 1328{,}5$

$$\pi(2x)-\pi(x) \ge \frac{x}{\ln x}-\frac{0{,}7x}{\ln^2 x}.$$

5. Conjecture d'Hardy-Littlewood.

CONJECTURE 1 (HARDY-LITTLEWOOD). *Pour tout* $x \ge 2$ *et pour tout* $y \ge 2$,

$$\pi(x+y) \le \pi(x)+\pi(y)?$$

THÉORÈME 6. *Pour tout* x *et pour tout* y *tel que* $x \le y \le \frac{7}{5}x\ln x\ln_2 x$, *on a*

$$\pi(x+y) \le \pi(x)+\pi(y).$$

En particulier, nous avons

$$\pi(2x) < 2\pi(x)$$

pour $x \ge 3$.

6. Dans les progressions arithmétiques.

DÉFINITION 1. *Soit*

$$\theta(x;k,l)=\sum_{\substack{p\le x\\ p\equiv l \bmod k}}\ln p,\quad \psi(x;k,l)=\sum_{\substack{p,\nu\\ p^\nu\le x,\ p\equiv l \bmod k}}\ln p$$

les fonctions de Chebyshev dans les progressions arithmétiques. Soit

$$\pi(x;k,l)=\sum_{\substack{p\le x\\ p\equiv l \bmod k}}1$$

la fonction qui compte le nombre de premiers inférieurs à x *et congrus à* l *modulo* k.

Pour un réel positif H donné, on dira que GRH(k,H) est satisfaite si, pour chaque χ modulo k, tous les zéros $\rho=\beta+i\gamma$, non triviaux, de la fonction de Dirichlet $L(s,\chi)$ tels que $|\gamma| \le H$ vérifient $\beta = 1/2$.

Étant donnés χ un caractère de conducteur k, $H \ge 1000$ et $\rho=\beta+i\gamma$ un zéro de $L(s,\chi)$ avec $|\gamma| \ge H$ alors il existe une constante $C_1(\chi,H)$ calculable (voir [7]) telle que

$$1-\beta \ge \frac{1}{R\ln\bigl(k|\gamma|/C_1(\chi,H)\bigr)}.$$

THÉORÈME 7. *Soit un entier* $k \geq 1$. *Soit* $R = 9{,}645908801$. *Soit* $H \geq 1000$. *Supposons GRH(k,H). Soit* $C_1(k)$ *défini par*

$$C_1(k) = \min_{\chi \bmod k} C_1(\chi, H_\chi).$$

Soit X_0, X_1, X_2 *et* X_3 *tels que*

$$\frac{e^{X_0}}{\sqrt{X_0}} = H\sqrt{\frac{k\varphi(k)}{2\pi C_1(k)}}, \quad \frac{e^{X_1}}{X_1} = 10\varphi(k),$$

$$X_2 = kC_1(k)/\big(2\pi\varphi(k)\big), \quad X_3 = \frac{2k\pi e}{C_1(k)\varphi(k)}.$$

Soit $C = \min\big(C_1(k), 32\pi\big)$ *et* $X_4 = \max(10, X_0, X_1, X_2, X_3)$.

Posons

$$\varepsilon(X) = 3\sqrt{\frac{k}{\varphi(k)C}}X^{1/2}\exp(-X).$$

Alors pour tout réel x *tel que* $X = \sqrt{\frac{\ln x}{R}} \geq X_4$ *nous avons*

$$|\psi(x;k,l) - x/\varphi(k)|, \quad |\theta(x;k,l) - x/\varphi(k)| < \varepsilon\left(\sqrt{\frac{\ln x}{R}}\right)x.$$

PROPOSITION 4. *Pour* $x > 1$ *et* $l = 1$ *ou 2,*

$$|\theta(x;3,l) - x/2| < 0{,}262\frac{x}{\ln x}.$$

THÉORÈME 8. *Pour* $l = 1$ *ou 2,*

(i) $\frac{x}{2\ln x} < \pi(x;3,l)$ *pour* $x \geq 151$,

(ii) $\pi(x;3,l) < 0{,}55\frac{x}{\ln x}$ *pour* $x \geq 229869$.

Nous redémontrons de cette façon que

$$\frac{x}{\ln x} < \pi(x).$$

REFERENCES

1. R. P. Brent, *Irregularities in the Distribution of Primes and Twin Primes.* Math. Comp. (129) **29**(1975), 43–56.
2. E. Cesaro, *Sur une formule empirique de M. Pervouchine.* Comptes rendus hebdo. des séances de l'académie des sciences **CXIX**(1895), 848–849.
3. M. Cipolla, *La determinazione assintotica dell'n^{imo} numero primo.* Matematiche Napoli **3**(1902), 132–166.
4. N. Costa Pereira, *Estimates for the Chebyshev Function* $\psi(x) - \theta(x)$. Math. Comp. (169) **44**(1985), 211–221.
5. Jean-Pierre Massias et Guy Robin, *Bornes effectives pour certaines fonctions concernant les nombres premiers.* Publication Département Math. de Limoges et Journal Th. Nombres de Bordeaux **8**(1996), 213–238.

6. O. Ramaré, *Short effective intervals containing primes.* Journal of Number Theory, à paraître.
7. O. Ramaré et R. Rumely, *Primes in arithmetic progressions.* Math. Comp. (213) **65**(1996), 397–425.
8. Hans Riesel, *Prime Numbers and Computer Methods for Factorization.* Birkhäuser, 1985.
9. Guy Robin, *Estimation de la fonction de Tchebychef θ sur le $k^{\text{ième}}$ nombre premier et grandes valeurs de la fonctions $\omega(n)$, nombre de diviseurs premiers de n.* Acta Arith. (4) **42**(1983), 367–389.
10. J. Barkley Rosser, *The n-th prime is greater than $n \log n$.* Proc. London Math. Soc. (2) **45**(1939), 21–44.
11. J. Barkley Rosser, *Explicit Bounds for some functions of prime numbers.* Amer. J. Math. **63**(1941), 211–232.
12. J. Barkley Rosser et L. Schoenfeld, *Approximate Formulas for Some Functions of Prime Numbers.* Illinois J. Math. **6**(1962), 64–94.
13. J. Barkley Rosser et L. Schoenfeld, *Sharper Bounds for the Chebyshev Functions $\theta(x)$ and $\psi(x)$.* Math. Comp. (129) **29**(1975), 243–269.
14. L. Schoenfeld, *Sharper Bounds for the Chebyshev Functions $\theta(x)$ and $\psi(x)$, II.* Math. Comp. (134) **30**(1976), 337–360.
15. J. van de Lune, H. J. J. te Riele et D.T. Winter, *On the Zeros of the Riemann Zeta Function in the Critical Strip, IV.* Math. Comp. (174) **46**(1986), 667–681.

Département de Mathématique
LACO, UPRES A 6090
123 avenue Albert Thomas
87060 LIMOGES cedex
France
email: dusart@unilim.fr

C. R. Math. Rep. Acad. Sci. Canada Vol. **21** (2), 1999 pp. 60–64

HAMILTONIAN MECHANICS ON PRINCIPAL BUNDLES

RICHARD CUSHMAN AND JĘDRZEJ ŚNIATYCKI

Presented by J. E. Marsden, FRSC

ABSTRACT. We discuss reduction for Hamiltonian systems with a symmetry group which acts properly and freely on configuration space.

RÉSUMÉ. Nous étudions la réduction pour les systèmes hamiltoniens dont le groupe de symétrie agit librement et proprement sur l'espace de configuration.

Here we consider a Hamiltonian system with configuration space Q, and Hamiltonian of the form $H = K + \pi_Q^* V$, where $K: T^*Q \to \mathbf{R}$ is the kinetic energy defined by a Riemannian metric k on Q, $V: Q \to \mathbf{R}$ is the potential energy, and $\pi_Q: T^*Q \to Q$ is the cotangent bundle projection. We assume that $\pi: Q \to M$ is a principal fibre bundle with structure group G and that the action of G lifted to T^*Q is a symmetry of our Hamiltonian H. This implies that the kinetic and potential energy are G-invariant. Therefore the metric k on Q is G-invariant. The distribution $\operatorname{hor} TQ$ on Q perpendicular to the fibres of $\pi: Q \to M$ with respect to the metric k defines a *principal connection* on Q. Restricting k to $\operatorname{hor} TQ$ gives rise to a metric g on M because $T_q\pi | \operatorname{hor}_q Q$ is an isomorphism onto $T_{\pi(q)}M$. The principal connection plays the fundamental role in the investigation of the structure of the dynamical system under consideration. In particular, it gives rise to connections and operators of covariant differentiation of sections of bundles associated to Q.

Let ω_Q be the canonical symplectic form on T^*Q. As the Hamiltonian vector field X_H on (T^*Q, ω_Q), corresponding to the Hamiltonian H, is G-invariant. Thus it projects to a vector field on the space $(T^*Q)/G$ of orbits of the lifted action of G on T^*Q. Let $\rho: T^*Q \to (T^*Q)/G$ be the G-orbit map. The *reduced space* $(T^*Q)/G$ is a fibre bundle over the *reduced configuration space* M, which splits as the direct sum of the coadjoint bundle $Q[\mathcal{G}^*]$ and the cotangent bundle T^*M over M, see [2]. In other words,

$$(T^*Q)/G = Q[\mathcal{G}^*] \oplus_M T^*M.$$

The action of the group $\operatorname{Aut}(Q)$ lifted to an action on T^*Q is Hamiltonian. For each X in the Lie algebra $\operatorname{aut}(Q)$ of $\operatorname{Aut}(Q)$, the action of the one parameter group

Received by the editors 5 February, 1999.
AMS subject classification: 58F05, 70H33.
Key words and phrases: Hamiltonian systems, reduction, symmetries.

$\exp tX$ on T^*Q is given by the flow of the Hamiltonian vector field of a function $\mathcal{J}_X$, which we call the *gauge momentum associated to* X. Since the Lie algebra $\mathrm{aut}(Q)$ is isomorphic to the space of sections of the adjoint bundle $Q[\mathcal{G}] \to M$, it follows that its dual $\mathrm{aut}(Q)^*$ is a space of distributions. The momentum map $\mathcal{J}: T^*Q \to \mathrm{aut}(Q)^*$ for the action of $\mathrm{Aut}(Q)$ on T^*Q is defined by $\mathcal{J}_X = \langle \mathcal{J} \mid X \rangle$. Here the pairing $\langle \mid \rangle$ is to be understood in the sense of distributions.

Let Γ be the map from T^*Q to the coadjoint bundle $Q[\mathcal{G}^*]$ formed from the composition of the projection $T^*Q \to \mathrm{ver}\, T^*Q$ and reduction map $\mathrm{ver}\, T^*Q \to (\mathrm{ver}\, T^*Q)/G$ followed by the bundle isomorphism of $(\mathrm{ver}\, T^*Q)/G$ onto $Q[\mathcal{G}^*]$. For each X in $\mathrm{aut}(Q)$, $\mathcal{J}_X = \langle \Gamma \mid \zeta_X \rangle$, where ζ_X is a section of the adjoint bundle $Q[\mathcal{G}] \to M$ corresponding to X and the pairing $\langle \mid \rangle$ is taken pointwise. We call Γ the *gauge momentum map* for the action of $\mathrm{Aut}(Q)$ on T^*Q.

Let $\Delta: T^*Q \to T^*M$ be the map formed from composition of the projection $T^*Q \to \mathrm{hor}\, T^*Q$ and reduction map $\mathrm{hor}\, T^*Q \to (\mathrm{hor}\, T^*Q)/G$ followed by the bundle projection of $(\mathrm{hor}\, T^*Q)/G$ on T^*M.

The G-invariant metric $\mathbb{k}^*$ on the fibres of T^*Q dual to $\mathbb{k}$ gives rise to a metric κ^* on the fibres of the coadjoint bundle $Q[\mathcal{G}^*] \to M: \alpha_m \to m$. Since the potential energy V on Q is G-invariant, it is the pull back of a function $\bar{V}$ on M under the projection map $\pi: Q \to M$. Let H_Γ be the function on $Q[\mathcal{G}^*]$ defined by

$$H_\Gamma(\alpha_m) = \tfrac{1}{2}\kappa^*(m)(\alpha, \alpha)$$

for every $\alpha_m \in Q[\mathcal{G}^*]$, and let H_Δ be the function on T^*M defined by

$$H_\Delta(y) = \tfrac{1}{2}g^*(x)(y, y) + \pi_M^* \bar{V}(y)$$

for every $y \in T_x^*M$. The Hamiltonian H on T^*Q admits a decomposition

$$H = \Gamma^* H_\Gamma + \Delta^* H_\Delta.$$

Let $t \mapsto p(t)$ be an integral curve the Hamiltonian vector field X_H on T^*Q. Its projection on the reduced space $(T^*Q)/G$ is uniquely determined by its projections $t \mapsto \gamma(t) = \Gamma\big(p(t)\big)$ on $Q[\mathcal{G}^*]$ and $t \mapsto y(t) = \Delta\big(p(t)\big)$ on T^*M, respectively. Let $t \mapsto x(t)$ be the image of $t \mapsto y(t)$ under the cotangent bundle projection $\pi_M: T^*M \to M$ and let σ be a section of the coadjoint bundle $Q[\mathcal{G}^*] \to M$ over the curve $t \to x(t)$ such that $\gamma(t) = \sigma\big(x(t)\big)$. We now describe the differential equations satisfied by the curves $t \mapsto \gamma(t)$ and $t \mapsto y(t)$.

On the one hand, the curve $t \mapsto \gamma(t)$ on $Q[\mathcal{G}^*]$ satisfies the nonautonomous *gauge momentum balance equation*

$$\frac{d\mathcal{J}_X}{dt} = -\tilde{X} \lrcorner dH, \quad \text{for all } X \in \mathrm{aut}(Q). \tag{1}$$

Here $\tilde{X}$ is the natural lift of X to T^*Q. Equation (1) can be decomposed into a nonautonomous *velocity equation*

$$(X_H \lrcorner \omega_Q - dH) \mid (\ker T\pi \cap \ker T\pi_Q) = 0 \tag{2}$$

and an autonomous *gauge momentum equation* for the section σ

$$\nabla_{\dot{x}}\sigma = T\Gamma\big(\omega_Q^\flat(\Gamma^* \operatorname{ver} dH_\Gamma)\big). \tag{3}$$

Here $\operatorname{ver} dH_\Gamma$ is the vertical part of the 1-form dH_Γ on $Q[\mathcal{G}^*]$ with respect to the connection on $Q[\mathcal{G}^*]$ associated to the principal connection. In addition, $\omega_Q^\flat: T^*Q \to TQ$ is the bundle isomorphism induced by the canonical symplectic form ω_Q. Thus the right hand side of (3) is a vector field on $Q[\mathcal{G}^*]$ along $\sigma(t)$. If the velocity equation is satisfied, the gauge momentum equation is equivalent to the conservation of the *principal momentum* map $J: T^*Q \to \mathcal{G}^*$ corresponding to the natural lift to T^*Q of the principal action of G on Q.

On the other hand, the curve $t \mapsto y(t)$ satisfies a nonautonomous Hamiltonian differential equation on T^*M, which depends on the lift of the curve $t \mapsto x(t)$ on the base to a solution $\sigma\big(x(t)\big)$ of the gauge momentum equation. Explicitly,

$$\dot{y} \lrcorner \omega_M = dH_\Delta - \dot{y} \lrcorner \pi_M^* \langle \sigma \mid \widehat{\Omega} \rangle + \Theta. \tag{4}$$

Here ω_M is a canonical symplectic form on the cotangent bundle $\pi_M: T^*M \to M$ of M, $\widehat{\Omega}$ is the 2-form on M with values in the fibres of the adjoint bundle $Q[\mathcal{G}] \to M$ induced from the curvature 2-form Ω of the principal connection $\operatorname{hor} TQ$, and $\Theta = \Delta_*(\Gamma^* \operatorname{hor} dH_\Gamma)$.

Assume now that our G-invariant Hamiltonian system is acted on by additional constraining forces such that the resulting motions have velocities in a G-invariant distribution D on Q. These additional forces are not specified. However, it is assumed that that the work they do on virtual displacements in D vanishes. Such constraining forces are called *perfect*. Furthermore, assume that the metric k on Q is Riemannian. The Legendre transformation for our system is given by $\mathcal{L}: TQ \to T^*Q: v \mapsto \mathrm{k}^\sharp(v)$. The requirement that the velocities of our system should be in D implies that the curves on T^*Q describing the motion of the constrained system have to lie in the *constraint manifold* $N = \mathcal{L}(D) = \{\mathrm{k}^\sharp(v) \in T^*Q \mid v \in D\}$.

Let F be the distribution on T^*Q, which is the pull back of D by the cotangent bundle projection $\pi_Q: T^*Q \to Q$, that is, $F = \{u \in T(T^*Q) \mid T\pi_Q(u) \in D\}$. In [3] Marle has shown that, under the assumptions made here, the following direct sum decompositions hold

$$(TN)^{\omega_Q} \oplus F = T_N TQ \quad \text{and} \quad TN \oplus F^{\omega_Q} = T_N TQ. \tag{5}$$

The vector field Y on $N \subseteq T^*Q$ describing the constrained motion satisfies the equation

$$(Y \lrcorner \omega_Q)_F = d_F H. \tag{6}$$

In the above equation the subscript F denotes the F-component of a 1-form with respect to the decomposition $(TN)^{\omega_Q} \oplus F = T_N TQ$. Equation (6) and the condition that Y is a vector field on N determine Y uniquely, see also [4].

Let $\bar{D}$ be the projection of D under $T\pi$. Then the image of N under the map $\Delta: T^*Q \to T^*M$ is $\{g^\sharp(u) \in T^*M \mid u \in \bar{D}\}$. Moreover, the image of $N \cap \operatorname{ver} T^*Q$ under the map $\Gamma: T^*Q \to Q[\mathcal{G}^*]$ is a subbundle $S = \Gamma(N \cap \operatorname{ver} T^*Q)$ of $Q[\mathcal{G}^*]$. Finally, observe that we have the direct sum decomposition

$$D = (D \cap \operatorname{ver} TQ) \oplus \big((D \cap \operatorname{ver} TQ)^\perp \cap D\big),$$

where the superscript $\perp$ denotes the orthogonal complement with respect to the Riemannian metric on Q. Since $T\pi\big((D \cap \operatorname{ver} TQ)^\perp \cap D\big) = \bar{D}$, for each $q \in Q$ there exists a map $\Lambda_q: \bar{D}_{\pi(q)} \to (D_q \cap \operatorname{ver} T_qQ)^\perp \cap D$ such that

$$\operatorname{lift}_q T\pi(u) + \Lambda_q T\pi(u) = u \quad \text{for all } u \in (D_q \cap \operatorname{ver} T_qQ)^\perp \cap D.$$

Here $\operatorname{lift}_q: T_{\pi(q)}M \to T_qQ$ is the horizontal lift with respect to the principal connection. The map Λ induces a map

$$\Sigma: \Delta(N) \to Q[G^*]: y \to \Sigma(y) = \Gamma\Big(k^\sharp\big(\Lambda_{q_x}(g^\flat(y))\big)\Big).$$

It follows that the image of N under the map Γ is $S \oplus_M \operatorname{range} \Sigma$ and the image of N under the G-orbit map ρ is $S \oplus_M \operatorname{graph} \Sigma$.

Let $\operatorname{ver}(D)$ denote the projection of D on the vertical distribution $\operatorname{ver} TQ$. We have

$$\operatorname{ver}(D) = (D \cap \operatorname{ver} TQ) \oplus_M \operatorname{range} \Lambda.$$

We have the following characterization of the gauge momenta on the constraint manifold N. For every $X \in \operatorname{aut}(Q) \cap \operatorname{ver}(D)^\perp$,

$$\mathcal{J}_X(p) = 0, \quad \text{for all } p \in N.$$

For each vector field $\check{X}$ on M with values in D, $\Lambda\check{X} \in \operatorname{aut}(Q)$ and

$$\mathcal{J}_{\check{X}}(p) = \mathrm{k}\Big(\Lambda_q\big(g^\flat \Delta(p)\big), \Lambda_q \check{X}\big(\pi(q)\big)\Big), \quad \text{for all } p \in N_q.$$

The values of the gauge momenta $\mathcal{J}_X$ corresponding to $X \in \operatorname{aut}(Q)$ with values in D are unrestricted on the constraint manifold N.

As in [1] the reduced equations of motion of the constrained system split into the momentum equation and the reduced equation. The *momentum equation* is

$$\frac{d\mathcal{J}_X}{dt} = -\big\langle dH_\Gamma\big(\Gamma(p)\big) \mid \Gamma_* \check{X}\big(\Gamma(p)\big)\big\rangle - \mathrm{g}^*(x)\big(\langle \Gamma(p) \mid \nabla \zeta_X(x)\rangle, \Delta(p)\big)$$

for all $X \in \operatorname{aut}(Q)$ with values in D. The restriction of the gauge momentum map Γ to $X \in \operatorname{aut}(Q)$ with values in D is the *nonholonomic momentum map*

of [1]. The *reduced equation* is

$$\begin{aligned}
&\omega_M\Big(T\Delta(Y(p)), Z(p)\Big) \\
&\quad = -\Big\langle \Gamma(p) \,\Big|\, \widehat{\Omega}\Big(T\pi_M T\Delta(Y(p)), (\pi_M)_* Z\Big(\pi_M(\Delta(p))\Big)\Big)\Big\rangle \\
&\qquad + \langle dH_\Delta \mid Z(\Delta(p))\rangle + \mathrm{g}^*(x)\big(\langle \Gamma(p) \mid \nabla_x \zeta_{\Lambda(\pi_M)_* Z}(x)\rangle, \Delta(p)\big) \\
&\qquad + \frac{d\mathcal{J}_{\Lambda(\pi_M)_* Z}}{dt}(p) + \big\langle dH_\Gamma(\Gamma(p)) \,\big|\, \Gamma_* \Lambda(\pi_M)_* Z(\Gamma(p))\big\rangle \\
&\qquad + (\nabla_{(\pi_M)_* Z} H_\Gamma)(\Gamma(p))
\end{aligned}$$

for every vector field Z on T^*M which projects to a vector field $(\pi_M)_* Z$ on M with values in $\bar{D}$.

The authors are greatly indebted to Jerry Marsden for his interest in this work and his helpful comments.

REFERENCES

1. A. Bloch, P. S. Krishnaprasad, J. E. Marsden and R. M. Murray, *Nonholonomic mechanical systems with symmetry.* Arch. Rat. Mech. Anal. **136**(1996), 21–99.
2. H. Cendra, D. D. Holm, J. E. Marsden and T. S. Ratiu, *Lagrangian Reduction, the Euler–Poincaré Equations, and Semidirect Products.* Trans. Amer. Math. Soc. **186**(1998), 1–25.
3. C.-M. Marle, *Reduction of constrained mechanical systems and stability of relative equilibria.* Comm. Math. Phys. **174**(1995), 295–318.
4. J. Śniatycki, *Non-holonomic Noether Theorem and Reduction of Symmetries.* In: Pacific Institute of Mathematical Sciences Workshop on Nonholonomic Constraints in Dynamics (Eds. R. Cushman and J. Śniatycki), Reports on Mathematical Physics **41/2**(1998), 5–23.

Mathematics Institute
University of Utrecht
Budapestlaan 6
3508TA Utrecht
The Netherlands
email: cushman@math.uu.nl

Department of Mathematics and Statistics
University of Calgary
Calgary, Alberta
T2N 1N4
email: sniat@math.ucalgary.ca